The Age of the Universe and the Mass of Stars by Gravitational Lensing

---

Dissertation

zur

Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Jonathan Coles

aus Grossbritanien

Promotionskomitee

Prof. Dr. George Lake (Vorsitz)
Dr. Prasenjit Saha

Zürich 2010

Contents

Acknowledgments

I won't pretend to be able to capture in just a few sentences the contributions of so many people to my personal and professional life over the last five years. I only hope to give a flavor, a hint, of the relationships and influences of my friends, family, and coworkers. I would therefore like to truly thank the following people for the rich and varied experiences I have had since coming to Zurich:

- Prasenjit Saha for his continuous help, encouragement, and ideas, without which I may still be working on this thesis.
- George Lake for taking me on as his student with hardly any physics experience and then letting me run free with my copious projects and interests.
- Ben Moore for his post to Slashdot that attracted my attention to Zurich, and for the countless hours spent in the Nelson discussing the nature of the Universe, its origins, and its future.
- My parents and sister for always being there to talk through my problems and finally pushing me to finish my thesis in time for their visit.
- Rebekka Bieri for her constant support, love, and encouragement in the last few months.
- Justin Read and Oscar Agertz. Two great friends who really made Zurich feel like home.
- Joachim Stadel for endless discussions about computers and science.
- Michèle Erat for her love and friendship through these crazy years.
- Doug Potter, Lea Giordano, Simone Callegari, Hanni Lux, and Francis Froborg for providing enough distractions through games and other adventures that I retained my sanity and didn't work *too* hard.
- Wyn Evans for agreeing on very short notice to be my external reader.
- All of my other colleagues who makes working at the University such a pleasure.

There are naturally many, many other people that I would love to thank. If you are not mentioned above please know that I still owe you a great deal of appreciation for all the kindness and excitement that you brought into my life.

Summary

In this thesis I review the current understanding of cosmology and how it relates to the field of gravitational lensing. One important cosmological parameter is the value of the Hubble constant H_0, which measures the expansion rate of the Universe today and also acts as a scaling factor. An accurate value of H_0 is critical for understanding not only the history of the Universe but its future as well. One can show that the age of the Universe is proportional to H_0^{-1}. Three papers of this thesis are dedicated to modeling gravitational lenses in order to measure H_0^{-1}. The most accurate models have shown that the Universe is $13.7^{+1.8}_{-1.0}$ Gyr old. In a further application, I explore how one can estimate the mass of a star in our galaxy by considering the microlensing effects on the light of a background galaxy. The findings suggest that low mass stars and brown dwarfs may be weighed to within 5% accuracy.

Zusammenfassung

In dieser Doktorarbeit wurde das derzeitige Verständnis der Kosmologie und ihrem Zusammenhang mit Gravitationslinsen betrachtet. Ein wichtiger kosmologischer Parameter ist der Wert der Hubble Konstante H_0 welche die heutige Expansionsrate des Universums misst und zudem als Skalarfaktor dient. Ein präziser Wert von H_0 ist für das Verständnis der Geschichte des Universums wie auch dessen Zukuft fundamental. Es kann gezeigt werden, dass das Alter des Universums direkt proportional zu H_0^{-1} ist. Die drei Paper dieser Doktorarbeit sind der Berechnung von H_0^{-1} mit Hilfe der Gravitationslinsen gewidmet. Das genauste Model berechnet das Alter des Universums auf $13.7^{+1.8}_{-1.0}$ Gyr. Als weitere Anwendung wurde der Mikrolinseneffekt zur Berechnung der Masse eines Sterns in unsere Galaxie bei Betrachtung des Lichtes einer Hintergrund Galaxie benutzt. Die Resultate zeigen eine fünfprozentige Genauigkeit bei der Berechnung von leichten Sternen sowie braunen Zwergen auf.

1 The Universe Through a Lens

1.1 Cosmology in a Nutshell

Our current picture of the Universe is a humbling one. We live on a planet orbiting an average star on the outskirts of a spiral galaxy that contains about 400 billion other stars. The Galaxy itself is only one of about 100 billion other observable galaxies. The Universe is expanding, with more distant galaxies receding from us at an accelerating rate. After billions and billions of years all of the surrounding galaxies will have disappeared beyond the horizon of the visible Universe, speeding away on the fabric of space faster than the speed of light. Should humanity survive this long, we will be alone; cut off completely from the rest of the Universe. If a new intelligent species were to evolve on another planet, their view of the Universe would be radically different from ours. There would be no evidence of other galaxies or of the expansion. As such, we live at a very favorable time, about 14 billion years since the Big Bang.

Much of our theoretical understanding of the Universe began with Einstein in 1915 and his theory of General Relativity (GR) (Einstein, 1916). The Einstein Field Equations

$$R_{\alpha\beta} - \frac{1}{2}Rg_{\alpha\beta} = \frac{8\pi G}{c^4}T_{\alpha\beta} \tag{1.1}$$

relate the geometry of space on the left side to the energy density of space of the right side. The Ricci tensor $R_{\alpha\beta}$ is includes second derivatives of the metric $g_{\alpha\beta}$ and in qualitative terms is similar to the Poisson equation

$$\nabla^2\Phi = 4\pi G\rho \tag{1.2}$$

used in Newtonian dynamics. Solving the field equations equates to finding an appropriate metric. One familiar metric describing flat space is the Euclidean metric $ds^2 = dx^2 + dy^2 + dz^2$. If the energy tensor $T_{\alpha\beta}$ takes a non-trivial form, however, the metric will become complicated.

The early success of GR was due to three things. First, it explained a correction to the precession of Mercury's orbit. All of the other planets in the system orbit the sun in a manner consistent with Kepler's laws of planetary motion and the laws of motion derived by Newton. Mercury's orbit, however, does not remain on a closed ellipse, but precesses, creating a rosette pattern. The effect is largely due to

Jupiter but there were some unexplained discrepancies. The additional effect of the curvature of space due to the Sun predicted by GR provided the correction needed to perfectly describe the orbit.

Second, according to the theory of Special Relativity published ten years earlier (Einstein, 1905), a clock moving relative to an observer will appear to run more slowly than a clock moving with the observer. If the first clock is brought into the frame of the observer, necessarily accelerating it, then indeed that clock will show that it has drifted. The founding principle of GR is the **equivalence principle**, which states that acceleration and a gravitational field are indistinguishable. Two clocks placed at different altitudes, and hence slightly different gravitational fields, should also experience a drift. This has been experimentally confirmed and is known as gravitational time dilation. This effect is not to be ignored. The global positioning system is sensitive enough that the effects of special and general relativity must be corrected for!

Third, the first test of the predictive power of GR was performed in 1919 by Sir Arthur Eddington and his team (Eddington, 1919; Dyson et al., 1920) during a total solar eclipse. It was already known that light should be deflected by a massive body but GR predicted an amount twice as much as Newtonian dynamics. During the eclipse the position of a star near to the sun was observed shifted by precisely the amount predicted by GR. This was the first confirmed lensing event.

Even with the great success of GR, the equations presented a problem. They allowed for a Universe that could expand or contract. The predominant view at the time was that the Universe was finite, unchanging, and eternal. To correct for this, mostly philosophical, view a new term $\Lambda g_{\alpha\beta}$ was added to the field equations that exactly canceled out the expansion. These objections to expansion did not stop others from considering its consequences, however.

In 1922, Alexander Friedmann Friedman (1922), and then later in 1927, Georges Lemaître (Lemaître, 1927), demonstrated that by assuming the **cosmological principle**, one can find a solution to the field equations that describes the evolution of the Universe and that naturally includes an expansion factor. The cosmological principle states there is no preferred location or direction in the Universe on large scales. In other words, it does not matter where one stands, the Universe looks the same in all directions. By large scales we mean scales of order 200 Mpc. To appreciate the size, note that Andromeda, our closest, largest galactic neighbor is less than 1 Mpc away.

Although Friedmann died in 1925, Lemaître did not have to wait long to see his work confirmed. Edwin Hubble (Hubble, 1929) showed that distant galaxies are in fact receding from us with a redshift proportional to their distance. If one believes that the Earth is not at the center of the Universe, the implication is that the Universe is expanding everywhere isotropically. This was the death knell for Λ as

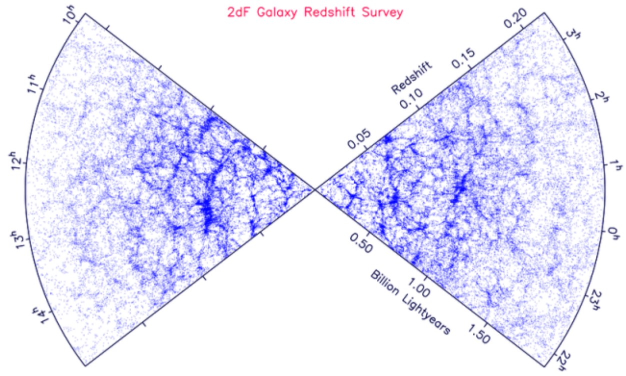


Figure 1.1: The distribution of galaxy as a function of distance and redshift from the 2dF survey. On large scales the Universe begins to look homogeneous.

it was understood in the early 20th century although the idea did take some time to become accepted. Even a year after Hubble's discovery Eddington (Eddington, 1930) wrote about Einstein's field equations, "On general philosophical grounds there can be little doubt that this form of the equations [with Λ] is correct rather than his earlier form."

Ironically, recent evidence has demonstrated that the growth of the Universe is actually *accelerating*. The Λ term has now been reintroduced, but as a place holder for an extra constituent of the Universe, on par with matter and radiation. Termed "dark energy," its origin and behavior have yet to be fully understood but recent observations with space satellites such as the Wilkinson Microwave Anisotropy Probe (WMAP) have shown that it is responsible for nearly 73% of the total energy density Ω of the Universe. All of the matter everywhere makes up nearly 27% and the only a marginal fraction is radiation. In terms of Ω, which is defined to be unity, the energy densities are

$$\Omega_m \sim 0.272 \quad \Omega_\gamma \sim 0 \quad \Omega_\Lambda \sim 0.728 \tag{1.3}$$

We now know from surveys such as 2dF Galaxy Redshift Survey (Peacock et al., 2001) that on very large scales the Universe can indeed be assumed to be isotropic and homogeneous. Fig. 1.1 shows the distribution of galaxies in the survey. The density appears to diminish with distance because the galaxies become increasingly fainter and harder to detect.

If one demands isotropy then the simplest line element for the metric $g_{\alpha\beta}$ is

$$ds^2 = -c^2dt^2 + a(t)^2[dr^2 + S_k(r)^2(d\theta^2 + \sin^2\theta d\phi^2)] \tag{1.4}$$

where

$$S_\kappa(r) = \begin{cases} R_0 \sin(r/R_0) & \text{if } \kappa = +1 \\ r & \text{if } \kappa = 0 \\ R_0 \sinh(r/R_0) & \text{if } \kappa = -1 \end{cases}, \tag{1.5}$$

R_0 is the radius of curvature of the Universe–a scaling factor–and r is the distance from an observer in comoving coordinates. This line element admits only three possible geometries that are consistent with the cosmological principle. That these are the only three was proven around 1935 by Howard Robertson and Arthur Walker (Robertson, 1935, 1936a,b; Walker, 1935). The metric has thus become known as the Robertson-Walker (RW) metric.

The particular sign of κ determines which geometry space will take. If $\kappa = +1$ space is like the surface of a sphere, if $\kappa = 0$ it is flat as in Euclidean geometry, and if $\kappa = -1$ it is shaped like a horse saddle. The presence of $a(t)$ is to allow for a general scaling that is only dependent on time, not on position.

The cosmological principle also implies that the energy tensor $T_{\alpha\beta}$ in the field equations is diagonal. This greatly simplifies the equations and with the substitution of the RW metric one can derive the Friedmann-Lemaître equation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{\kappa c^2}{R_0^2}\frac{1}{a^2} \ . \tag{1.6}$$

Solving for a determines the expansion rate of the Universe and thus provides the metric solution to the field equations. This is the same rate that Hubble was measuring, although his value was nearly seven times too large because of errors in his distance calculations. Nevertheless, we define the **Hubble parameter** as

$$H \equiv \frac{\dot{a}}{a} \ . \tag{1.7}$$

At the present age of the Universe t_0 we also define $H_0 = H(t_0)$ as the **Hubble constant** and H_0^{-1} as the **Hubble Time**.

The Hubble constant acts as a scaling factor and is very important to cosmology. One place it appears is in the Taylor expansion of a around t_0:

$$a(t) = a(t_0) + \dot{a}(t_0)(t - t_0) + \cdots \ . \tag{1.8}$$

Dividing through by $a(t_0) = 1$ we find

$$a(t) = 1 + H_0(t - t_0) + \cdots \ . \tag{1.9}$$

In order to determine a we need to have a good estimate for H_0.

Because of the expansion, the time and distance that light travels between two points depends on a. The look back time is

$$t = \int_a^1 \frac{da}{\dot{a}} \tag{1.10}$$

and the distance traveled is

$$r = \int_a^1 \frac{da}{\dot{a}a} \ . \tag{1.11}$$

An observer on Earth looking at a distant object will see the light from the object as it was in the past. The apparent and actual size of the object is then also dependent on a. This is known as the angular diameter distance

$$D = S_k a \ . \tag{1.12}$$

The Friedmann-Lemaître equation is fundamental to cosmology. By understanding the evolution of the expansion factor, we can explore the history and future of the Universe. If we consider the curvature to have an effect similar to a density we can write

$$\Omega = \Omega_m + \Omega_\gamma + \Omega_\Lambda = 1 \tag{1.13}$$

with $\Omega_k = 1 - \Omega$. In terms of Ω, Eq. 1.6 can be written

$$\left(\frac{H}{H_0}\right)^2 = \frac{\Omega_m}{a^3} + \frac{\Omega_\gamma}{a^4} + \Omega_\Lambda + \frac{\Omega_k}{a^2} \ . \tag{1.14}$$

With Eq. 1.14 and the definition of H one can immediately derive an expression for the age of the Universe,

$$t_0 = H_0^{-1} \int_0^1 a \left[a\Omega_m + \Omega_\gamma + a^4\Omega_\Lambda + a^2\Omega_k\right]^{-1/2} da \ . \tag{1.15}$$

Determining a precise value of H_0^{-1} is not easy. Hubble had used his published relation for the scaling of velocity with distance,

$$d = H_0 cz \tag{1.16}$$

to estimate $H_0 \approx 500$ km/s/Mpc. While the velocity of galaxies is relatively easy to determine from redshift measurements, distance is particularly uncertain. More recent surveys of Cepheid variable stars and Type Ia supernovae have shown that Hubble's value was too large by about seven times solely due to his inaccurate distance measurements.

An alternative approach is to use gravitational lensing. Ironically, Albert Einstein (Einstein, 1936) wrote that "there is no great chance of observing this phenomenon." With more modern instruments than were available in Einstein's time, gravitational lensing has now become a fundamental tool in cosmology. Lensing has been used to weigh galaxies, discover evidence for dark matter and dark energy, and more recently has been pivotal in the detection of exoplanets orbiting distant stars.

1.2 Gravitational Lensing

Gravitational lensing occurs when the path of a ray of light is deflected from a straight line by the presence of a massive object. The object curves space in accordance with the predictions of GR and the light ray follows this curvature. A

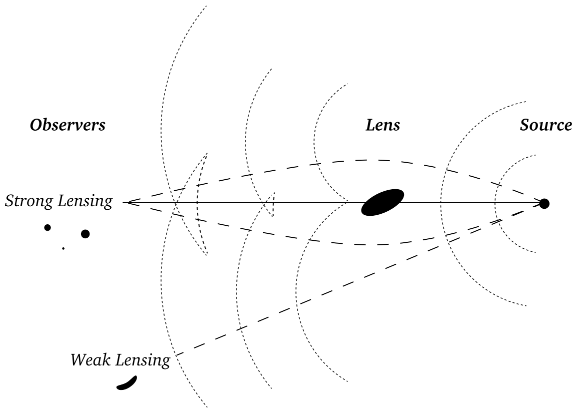


Figure 1.2: A simple lensing scenario. Light from the source is bent strongly in the presence of a galaxy creating multiple images for the observer. Further from the galaxy the image of the source is only distorted. These two regimes are known as strong and weak lensing, respectively.

simple picture is shown in Fig. 1.2. The observed effect is very similar to the lensing effect when light passes through glass. If the source, lens, and observer are very nearly aligned multiple images of the source will appear and those images may be magnified or demagnified. This is a regime called **strong lensing**. The number of images that appear depends on the exact alignment, but it will always be an odd number. Typically, though one image is highly demagnified to the point where it becomes unobservable. If the alignment is perfect and the lens has nearly spherical symmetry a ring of infinitely many images forms, called an Einstein ring.

Until 1979 lensing remained a purely theoretical curiosity and seemed to go out of favor until the 1960s when Sjur Refsdal revisited it. In 1964 he published a new derivation of the lens equations in the case where one star lenses the light of another star and that in fact the phenomenon should be observable (Refsdal, 1964a). In the same journal edition he also discussed the possibility of measuring H_0 (Refsdal, 1964b). Then two years later he discussed weighing stars by lensing (Refsdal, 1966).

The first lens to be found was 0957+561 by Weymann et al. (1979) and is now a typical example of a quasar that has been strongly lensed into two visible images. The lens is shown in Fig. 1.3. The two small points of light are from the lensed quasar while the large colored region is the foreground galaxy. If the alignment is not good enough to produce strong lensing, there will be only one image of the source but it will appear warped. This effect is known as **weak lensing**. A striking example is shown in Fig. 1.4. This is a picture of the galaxy cluster Abell 1689. Each of the long arcs of light is the distorted image of galaxies lying behind the main cluster. The third regime is known as **microlensing**. This occurs when the size of the lens is too small to be resolved by conventional telescopes and all that is

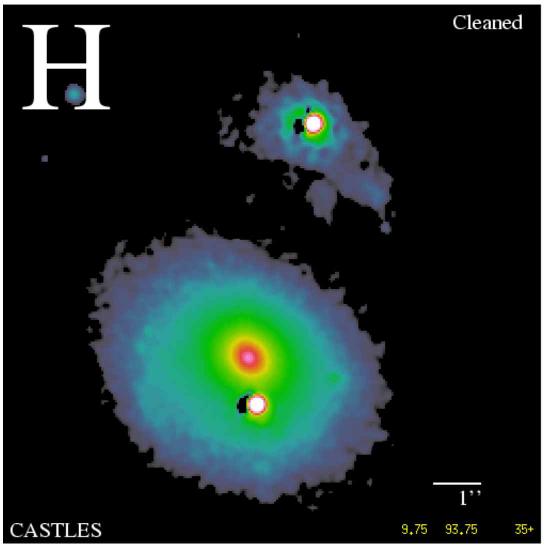


Figure 1.3: The first gravitational lens discovered, 0957+561. This is an example of a strongly lensed quasar by a foreground galaxy.

observed is a brightening of the source.

Fig. 1.5 depicts a gravitational lens schematically. The distance from the observer to the lens, the lens to the source, and the observer to the source are denoted by $D_{\rm OL}, D_{\rm LS}$ and $D_{\rm OS}$, respectively, and are usually angular diameter distances. It is typically assumed that the size of the lens is very small compared to $D_{\rm OL}$ and $D_{\rm OS}$ which allows us to assume that all of the mass of the lens is projected onto a plane at $D_{\rm OL}$.

The coordinate system is defined to be centered on the middle of the lens. In the absence of the lens the source would appear on the sky at an angular position $\boldsymbol{\beta}$. With the lens in place, the actual source position becomes unknowable. A light ray from the source travels to the lens plane, interacts with the lens at the impact parameter distance $\boldsymbol{b}$, and is deflected by the deflection angle $\boldsymbol{\alpha}$. When the ray arrives at the observer, it appears to be coming from an angular position $\boldsymbol{\theta}$.

Because of the thin lens approximation the time that the light ray interacts with the lens is so short that the Universe does not expanding significantly and the local background metric $\eta_{\alpha\beta}$ around the lens is flat, or Minkowskian. The only perturbations are due to the mass of the lens itself so we can write

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} \ . \tag{1.17}$$

where $h_{\alpha\beta} \ll 1$.

In the **weak field**, where one assumes a Newtonian gravitational potential $\Phi =$

Figure 1.4: Abell 1689 galaxy cluster. The massive gravitational field of the cluster lenses the light of background galaxies into long, beautiful arcs.

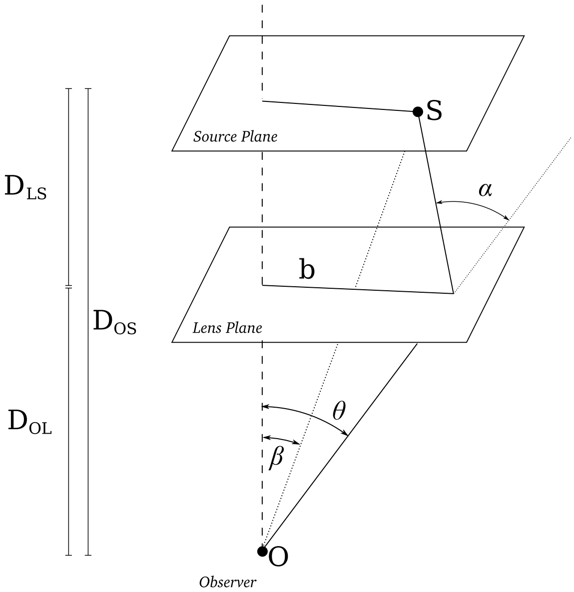


Figure 1.5: A schematic of a gravitational lens. In the absence of the lens, the source S would be observed at position $\boldsymbol{\beta}$, but with the lens the source appears at position $\boldsymbol{\theta}$. The original light ray impacted the lens plane a distance $\boldsymbol{b}$ from the origin and was deflected by an angle $\boldsymbol{\alpha}$.

$GM/r < 1$, the metric becomes

$$ds^2 = -\left(1 + \frac{2\Phi}{c^2}\right) c^2 dt^2 + \left(1 - \frac{2\Phi}{c^2}\right) dl^2 \ . \tag{1.18}$$

With this metric the bending angle for a point mass M is

$$\boldsymbol{\alpha} = \frac{4GM}{c^2} \frac{\boldsymbol{b}}{|\boldsymbol{b}|^2} \ . \tag{1.19}$$

If we treat the lens plane as a collection of point masses the total bending angle is just an integral over the plane,

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{4G}{c^2} \int_{\mathbb{R}^2} d^2\theta' \Sigma(\boldsymbol{\theta}') \frac{\boldsymbol{\theta} - \boldsymbol{\theta}'}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2} \tag{1.20}$$

where Σ is the projected mass density at a point.

A very useful quantity in lensing is the **time delay** $\varphi = t_l - t_d$, where t_d is the travel time a light ray would take from the source directly to the observer in the absence of the lens and t_l the time it takes with the lens in place. Any deviation from a straight line means that $t_l > t_d$. Since light always travels on null geodesics ($ds^2 = 0$) Eq. 1.18 becomes

$$c^2 dt^2 = \frac{1 - 2\Phi/c^2}{1 + 2\Phi/c^2} dl^2 \tag{1.21}$$

which can be approximated and integrated to find the travel time

$$ct = l - 2c^{-2} \int \Phi dl \ , \tag{1.22}$$

of which the first term is a geometric contribution and the second term, known as the Shapiro time delay, is the contribution of the gravitational potential. The second term is an inverse of the bending angle $\boldsymbol{\alpha}$ so that we write

$$ct = l - \nabla^{-2} \boldsymbol{\alpha} \ . \tag{1.23}$$

where we define ∇^{-2} to be an operator that solves Poisson's equation in two dimensions. In other words, if $\nabla^2 f(\boldsymbol{\theta}) = g(\boldsymbol{\theta})$ then $f(\boldsymbol{\theta}) = \nabla^{-2} g(\boldsymbol{\theta})$.

For the light ray traveling directly to the observer the Shapiro time delay vanishes and thus for φ we only need to consider the difference in path lengths for the lensed and unlensed rays. Under small angle approximations that difference is

$$\Delta l = \frac{D_{\mathrm{OL}} D_{\mathrm{OS}}}{2 D_{\mathrm{LS}}} (\boldsymbol{\theta} - \boldsymbol{\beta})^2 \tag{1.24}$$

In a cosmological context, where the D's are angular diameter distances, we must scale by $(1+z)$. We find then that the final time delay is

$$c\varphi = (1+z_L)\left\{\frac{D_{\mathrm{OL}}D_{\mathrm{OS}}}{2D_{\mathrm{LS}}}(\boldsymbol{\theta}-\boldsymbol{\beta})^2 - \nabla^{-2}\boldsymbol{\alpha}\right\} \ . \tag{1.25}$$

This form tends to be cumbersome, so we will rewrite it in dimensionless form. Let the dimensionless time delay τ be denoted

$$\tau \equiv \left[(1+z_L)\frac{D_{\mathrm{OL}}D_{\mathrm{OS}}}{D_{\mathrm{LS}}}\right]^{-1} c\varphi \ . \tag{1.26}$$

Next, define a dimensionless distance factor

$$c_{\mathrm{LS}} \equiv \frac{D_{\mathrm{LS}}}{D_{\mathrm{OS}}} \tag{1.27}$$

and a dimensionless surface density

$$\kappa \equiv \frac{4G}{c^2} D_{\mathrm{OL}}\Sigma \ . \tag{1.28}$$

This surface density is related to the critical density, which we define as

$$\Sigma_{\mathrm{crit}} \equiv \frac{c^2}{4GD_{\mathrm{OL}}} \ . \tag{1.29}$$

For a source at infinity, lensing will only occur when $\kappa \geq 1$. With these definitions we can write Eq. 1.25 as

$$\tau = \frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\beta})^2 - 2c_{\mathrm{LS}}\nabla^{-2}\boldsymbol{\alpha} \tag{1.30}$$

where $\boldsymbol{\alpha}$ is now also in dimensionless form

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \int_{\mathbb{R}^2} d^2\theta' \kappa(\boldsymbol{\theta}')\frac{\boldsymbol{\theta}-\boldsymbol{\theta}'}{|\boldsymbol{\theta}-\boldsymbol{\theta}'|^2} \ . \tag{1.31}$$

We can think of $\tau(\boldsymbol{\theta})$ as an arrival time surface. Fermat's principle tells us that images only appear at locations where $\nabla\tau(\boldsymbol{\theta}) = 0$. In other words, images appear at minimums, maximums, and saddle points of this surface. Under this restriction, we immediately derive from Eq. 1.30 the **lens equation**

$$\boldsymbol{\beta} = \boldsymbol{\theta} - 2c_{\mathrm{LS}}\boldsymbol{\alpha} \tag{1.32}$$

which maps images back onto the source plane.

The value of the arrival time tells us in which order the photons arrive. This feature will become critical when we consider applications of lensing. If the source varies in brightness, for instance, the change will appear first at the lowest minimum and last at the highest maximum. Measuring the time delay between successive changes at the images tells us about the shape of the arrival time surface.

Consider now the second derivative,

$$\nabla\nabla\tau = \mathbf{1} - 2c_{\mathrm{LS}}\nabla\boldsymbol{\alpha} \equiv \boldsymbol{K} \tag{1.33}$$

which is the Jacobian matrix that describes how small changes in image positions map back to the source plane. The inverse is called the **magnification tensor** $\boldsymbol{M} = \boldsymbol{K}^{-1}$. Under a suitable coordinate rotation $\boldsymbol{K}$ and $\boldsymbol{M}$ can be diagonalized, giving

$$\boldsymbol{K} = \begin{bmatrix} 1/e_1 & 0 \\ 0 & 1/e_2 \end{bmatrix} \qquad \boldsymbol{M} = \begin{bmatrix} e_1 & 0 \\ 0 & e_2 \end{bmatrix}, \tag{1.34}$$

where e_1, e_2 are the eigenvalues. The total magnification of the image is the determinant of $\boldsymbol{M}$, or in this case, just the product of the eigenvalues $|\boldsymbol{M}| = e_1 e_2$. If $e_1 = e_2$ then the image is simply magnified, but if $e_1 \neq e_2$ then the image is stretched along the eigenvectors. This stretching is called **shear** and its value it defined as

$$\mu = \frac{1}{2}(e_1^{-1} - e_2^{-2}) \ . \tag{1.35}$$

Up to now, it may be unclear how this can be put to any use. One particularly interesting and useful aspect of strong lensing occurs when the lensed object is a quasar. Many of the lensed quasars are variable; they change brightness with time, often on the order of days. Because the path length depends on the mass distribution the path lengths of two images will differ causing the fluctuating brightness to be observed in one image before the same fluctuation is observed in the other. This quantity can be measured and put to good affect.

Consider subtracting time delays of images at $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$:

$$\frac{\tau_1 - \tau_2}{\varphi_1 - \varphi_2} = c\left[(1+z_L)\frac{D_{\mathrm{OL}}D_{\mathrm{OS}}}{D_{\mathrm{LS}}}\right]^{-1} \propto H_0^{-1} \ . \tag{1.36}$$

The proportionality comes from each of the distances, which have a factor of H_0 in them. Since $(\varphi_1 - \varphi_2)$ is a measurable quantity, we need only to calculate the dimensionless time delays τ_1 and τ_2 in order to calculate H_0^{-1}. This is much easier said than done. The most significant problem with using lensing in this way, is that lensing is prone to degeneracies.

The simplest degeneracy is the **steepness degeneracy**. If we scale τ by a constant

$$\tau' = \lambda\tau \tag{1.37}$$

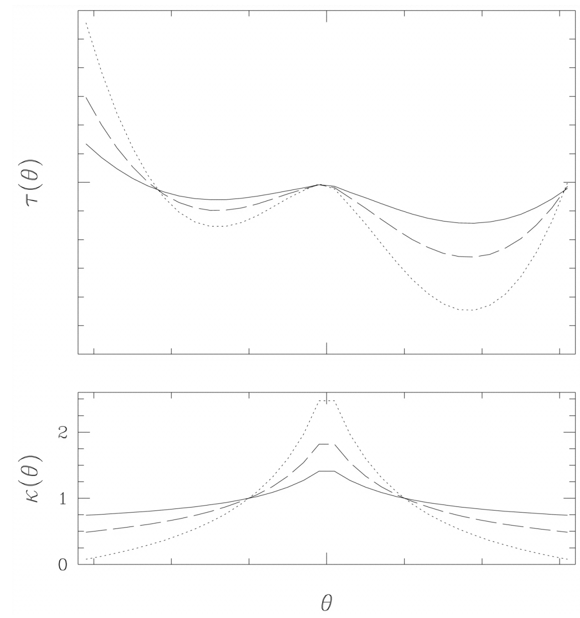


Figure 1.6: The steepness degeneracy illustrated. Scaling the arrival time surface $\boldsymbol{\tau}$ leaves the image positions (local minima and maximum) in the same places (*top*) while changing the steepness of the surface density κ (*bottom*). (Saha, 2000)

then we find that $\nabla\tau' = \nabla\tau$ so that the images appear in the same location. The reason for the name is that a constant scaling will make the the mass plane shallower or steeper for $\lambda < 1$ or $\lambda > 1$, respectively, as shown is Fig. 1.6.

One can imagine more complicated degeneracies too where $\tau = \lambda(\boldsymbol{\theta})\tau$. Although there is no known analytic form for $\lambda(\boldsymbol{\theta})$, other degeneracies have been explored in the literature. Shape degeneracies, for instance, twist the mass plane in complicated ways (Saha and Williams, 2006), while the surface density could also be represented as a sum of mass rings (Liesenborgs et al., 2008).

1.3 Applications of strong lensing

Despite the difficulties in breaking lensing degeneracies there is still much we can learn about the Universe through lensing. Weak lensing, often seen in galaxy clusters, provides a rich source of information about their mass distribution. Typically, many background galaxies are lensed, not just a single source, and so although the lensing is not very strong there is a wide range of samples to use to reconstruct plausible mass distributions. This is important for understanding the dynamics within those clusters. Some work has even explored using weak lensing to measure the equation of state of dark energy, the component of the Universe responsible for nearly 75% of the energy content and the accelerated expansion.

As we have already seen, strong lensing is particularly suited for studying galaxies and measuring H_0. Lensing can measure very accurately the enclosed mass inside the Einstein radius

$$\theta_E = \left(\frac{4GM}{c^2}\frac{D_{\mathrm{LS}}}{D_{\mathrm{OL}}D_{\mathrm{OS}}}\right)^{1/2} \tag{1.38}$$

This is generally near the position of the outermost observed image. In this way, lensing has provided strong evidence for the existence of dark matter. The discrepancy between the enclosed mass measured through lensing and the total mass measured from star light is significant enough to suggest the presence of a particle that interacts gravitationally but not electromagnetically.

Two papers of this thesis have explored measuring H_0 by strong lensing.

- *The Hubble Time Inferred from 10 Time Delay Lenses* (PAPER I)
- *A New Estimate of the Hubble Time with Improved Modeling of Gravitational Lenses* (PAPER II)

With one lens alone the range of physically possible H_0 values is large because of the degeneracies discussed above. In the related paper

- *COSMOGRAIL: The COSmological MOnitoring of GRAvItational Lenses VII: Time delays and the Hubble constant from WFI J2003-4723* (PAPER IV)

only a single lens was modeled. The errors in the estimate of H_0 are on the order of 20%. One can do much better, though, with multiple lenses by making the simple observation that the Hubble constant should be constant for all lenses. The technique employed in the first two papers combines data from multiple lenses with the additional constraint that for all lenses H_0 is the same. The results build on work first done by Saha and Williams (2004) which combined data from three lenses. The results are much more promising, with errors less than 10%.

The computer program PIXELENS is used to construct possible mass distributions that reproduce all of the observables. A solution for a particular lens

$$\boldsymbol{S} = (H_0^{-1}, \boldsymbol{\beta}, \gamma_1, \gamma_2, \kappa_1, \kappa_2, \ldots, \kappa_n) \tag{1.39}$$

consists of values for N cells, or pixels, that represent the density κ_n in a region of the lens on the sky. The lens equation is written in a discrete form

$$\tau(\boldsymbol{\theta}) = \frac{1}{2}(\boldsymbol{\theta} - \boldsymbol{\beta})^2 - 2c_{\mathrm{LS}} \sum_{n=1}^{N} \kappa_n Q_n + \gamma_1(\theta_x^2 - \theta_y^2) + \gamma_2 \theta_x \theta_y \tag{1.40}$$

where Q is the position dependent part of $\nabla^{-1}\boldsymbol{\alpha}$ evaluated over the pixel and γ_1, γ_2 are shear terms. The shear terms here are not related to the shear in Eq. 1.35. Their inclusion here is because we are only considering a finite area of the sky and it is possible that a neighboring galaxy may contribute to the potential in a weak fashion. In principle, one can add any number of shear terms to account for multiple objects.

The exact form of Q can be found in Paper II. The quantity we need for Eq. 1.36 is the difference in time delays $\Delta\tau = \tau_1 - \tau_2$ and since the source position is same for each τ, the β^2 term in Eq. 1.40 will disappear and $\Delta\tau$ will be linear in the unknowns.

If multiple lenses are used a solution, or model, is a set of sub-solutions with a shared H_0^{-1}

$$\boldsymbol{M} = \{\boldsymbol{S}_1, \boldsymbol{S}_2, \dots\} \qquad \boldsymbol{S}_1[H_0] = \boldsymbol{S}_2[H_0] = \dots \tag{1.41}$$

A model ensemble is a set of models

$$\boldsymbol{E} = \{\boldsymbol{M}_1, \boldsymbol{M}_2, \dots\} \tag{1.42}$$

where different models are allowed to have different H_0^{-1} values. PixeLens generates many different models using a Monte-Carlo Markov-Chain (MCMC) scheme.

To find a particular solution, PixeLens uses linear programming. A linear program consists of a linear *objective* function z subject to linear constraints. Normally one is interested in finding a solution that either minimizes or maximizes z. Without loss of generality we will assume that we wish to maximize z:

$$\begin{array}{ll} \text{maximize} & z = \boldsymbol{C}^T\boldsymbol{x} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \end{array} \tag{1.43}$$

where $\boldsymbol{x}$ is the solution matrix. The inequalities define a convex polygon called a simplex in an n-dimensional space, with all possible solutions lying inside. The solution to Eq. 1.43 will necessarily lie at a vertex of this polygon. Many algorithms have been developed to find such solutions, but one of the more successful is the **simplex algorithm**. The idea is to pick a vertex at random and walk to a neighboring vertex for which z is greater. The walking process continues until all neighboring vertices have a lower value.

PixeLens uses this algorithm in a unique way to explore the solution space. Since both $\Delta\tau$ and $\nabla\tau$ are linear in the unknowns we can use them to constrain a solution space consisting of models that perfectly satisfy the observables. Additional constraints called priors can also be added to avoid noisy or unphysical solutions. To begin, as with the simplex algorithm, a random vertex $\boldsymbol{v}_0$ is found. Let us also label this vertex $\boldsymbol{m}_0$, for reasons that will become clear. Then a *random* objective function is generated where each element of $\boldsymbol{C}^T$ is a number between $-1/2$ and $1/2$. The simplex algorithm maximizes this objective function, finding a solution at vertex $\boldsymbol{v}_1$. Imagine now extending a line from the current vertex $\boldsymbol{v}_n$ through the previous model $\boldsymbol{m}_{n-1}$ until the edge of the polygon is reached at a point $\boldsymbol{r}$. The new model is then chosen to be $\boldsymbol{m}_n = \boldsymbol{r} + q(\boldsymbol{v}_n - \boldsymbol{r})$ where $0 \leq q \leq 1$ is a random number. Repeating this process, a new random objective function is maximized, which leads to a new $\boldsymbol{v}_n$, which leads to a new model. Each model is added to the ensemble $\boldsymbol{E}$ except for $\boldsymbol{m}_0$. The result of this procedure is a collection of models that samples the solution space. Further analysis of this can be found in Paper II.

Having selected our models we can then investigate the distribution of H_0^{-1}. For this we turn to Bayesian statistics. The fundamental equations are known as the sum and product rules of probability

$$\mathrm{prob}(X|I) + \mathrm{prob}(\overline{X}|I) = 1 \tag{1.44}$$

and

$$\mathrm{prob}(X, Y|I) = \mathrm{prob}(X|Y, I) \times \mathrm{prob}(Y|I) \tag{1.45}$$

where X and Y are some propositions and I are our prior assumptions. From these two statements one can derive two important results: **Bayes' theorem** and **marginalization**. For clarity, we should think of Y as our hypothesis or model $\boldsymbol{M}$ and X as our observed data $\boldsymbol{D}$. In these terms Bayes' theorem is

$$\mathrm{prob}(\boldsymbol{M}|\boldsymbol{D}, I) \propto \mathrm{prob}(\boldsymbol{M}|\boldsymbol{D}, I) \times \mathrm{prob}(\boldsymbol{D}, I)\ . \tag{1.46}$$

and marginalization is

$$\mathrm{prob}(\boldsymbol{D}, I) = \int_{-\infty}^{+\infty} \mathrm{prob}(\boldsymbol{D}, \boldsymbol{M}) d\boldsymbol{M}\ . \tag{1.47}$$

Eq. 1.47 may appear strange because one is integrating over $d\boldsymbol{M}$. Let's instead write this in its discrete form, where we have N models. Then

$$\mathrm{prob}(\boldsymbol{D}|I) = \sum_{k=1}^{N} \mathrm{prob}(\boldsymbol{D}, \boldsymbol{M}_k|I) \tag{1.48}$$

The usefulness of marginalization comes from the fact that we can eliminate so-called nuisance parameters, parameters that are necessary for the problem but are in the end of no interest.

If we marginalize over the our ensemble $\boldsymbol{E}$ we find a distribution for H_0^{-1}. Adding more lenses, and thus more models, better constrains the distribution. The peak of the resulting curve is our "best guess."

With ten lenses it was found in Paper I that

$$H_0^{-1} = 13.5^{+2.5}_{-1.3}\ \mathrm{Gyr} \tag{1.49}$$

and in Paper II that

$$H_0^{-1} = 13.7^{+1.8}_{-1.0} \text{ Gyr} . \tag{1.50}$$

With additional lenses it is hoped to eventually reach the 1% level.

Unfortunately, a sufficient number of time-delay lenses with accurate measurements are not currently available. There are also technical challenges to overcome. Each lens that is added increases the number of variables and constraints in Eq. 1.43 by about N but the running time of the simplex algorithm is $O(N^3)$.

One of the advancements made in Paper I and Paper II was to parallelize the simplex algorithm. Without this improvement even ten or eleven lens would have been to great to run in any reasonable length of time.

Of course, one is only using the simplex algorithm to sample the interior of the solution space. If another algorithm were developed that accomplished this faster then the technical problems would be solved. All that is actually needed is another method to find random vertices of a simplex and then the model selection part of the above technique can still be used.

1.4 Applications of weak- and micro-lensing

Microlensing has typically been explored in the context of planet detection. The effect is very subtle and occurs when a star with planets happens to align with another star in the galaxy. During the alignment the light from the background star will microlensed and will be observed to grow in brightness. If the foreground star happens to be host to a planet then that planet may also act as a lens. One result (Beaulieu et al., 2006) from the Optical Gravitational Lens Experiment (OGLE) is shown in Fig. 1.7. The large curve is the magnification due to the star, while the small feature shown in the corner is due to an orbiting planet.

A novel application of microlensing was explored in Paper III, *Weak Microlensing* in which we used a statistical analysis to measure the mass of stars. We have shown that it may be possible to measure very low mass stars and brown dwarfs to within an accuracy of about 5%.

As was discussed early, lensing can measure the mass enclosed inside the Einstein radius θ_E very accurately. Unfortunately, this is only applicable if the observed image is well resolved. With microlensing the lens is too small and therefore only the brightening is observed. In the paper we prefer to discuss θ_E rather than the mass directly. This avoids unnecessary considerations about distance factors.

In Paper III we have considered how well we can measure the mass of a star in our galaxy that transits across a background galaxy. We know that the light from the background galaxy will be lensed by the star and that there will be a subtle brightening of the image which will, of course, depend on the mass. The

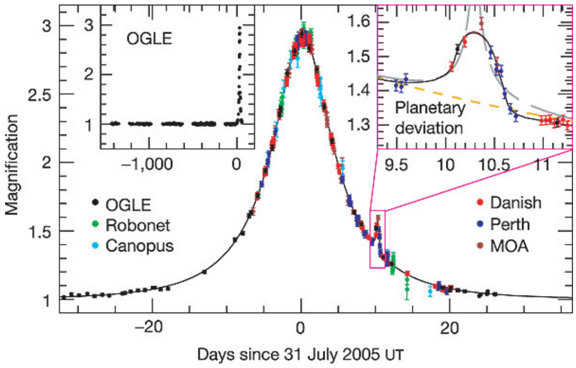


Figure 1.7: A signature of a planet from microlensing. An orbiting planet causes a slight increase in magnification of a background star in addition to the much larger effect of the foreground star.

brightening will be so small, however, as to be imperceptible to the human eye but the *computer* should be able to see the change.

The work to date is only theoretical and so the "observations" we use are computer generated. In the future we hope to apply the analysis to real data. The observations are based on a model for the light distribution of a galaxy. We used a modified $R^{1/4}$ model that incorporated Poisson noise and effects of a point spread function. This represented an unlensed observation S_0 before the star transit. We then assumed that a star with $\theta_E = \theta_{E,\mathrm{true}}$ would be observed at different time intervals across the galaxy and that the location of the star z was known. At each time t we lensed S_0 to produce a new observation S_t.

Since we in reality we do not known *a priori* the value of $\theta_{E,\mathrm{true}}$ our modeling procedure search for an estimate of θ_E that best reconstructs the observations.

For each estimate we use a least-squares analysis to fit test against the known observations. The θ_E that best models the observation is then deemed. To fully test the method we explored a large range of different parameters. We considered a variety of $\theta_{E,\mathrm{true}}$'s, photon counts from the observations, number of observations, and the distance of closest approach the star makes to the galaxy.

We found that in 79% of our 500 test cases we were able to recovery the true θ_E to within 11%. In 52% of the cases, to within 5%.

1.5 Focusing on the Future

New lenses are discovered on a regular basis now and future surveys promise to find hundreds if not thousands more. Although most of these new lenses will not have measured time delays the wealth of data will provide many opportunities for statistical analysis over a wide range of lensing galaxy types and morphologies. Some of these lenses will surely be the focus of further study and time delays will be forthcoming. As discussed earlier, with more time delay lenses we can hope to push the errors down on the estimate of H_0^{-1} to below 1%. This would place gravitational lensing into the regime of precision cosmology; becoming a very serious check on other such measurements.

In addition to just finding better estimates of H_0^{-1}, we have already seen other applications. Improvement in technology will allow us to use lensing on ever smaller scales. What is now only perceived as microlensing may one day be fully resolved.

Even today there are many avenues for further research. The nature of many lensing degeneracies is still not fully understood. Research already underway will explore this in more detail. It is also important to understand the limits of what can be achieved with current data so that we can think of new modeling techniques or other observations to perform. When lensing data is not enough, it is important to consider what other information is available that may come from a completely different field.

Bibliography

J.-P. Beaulieu, D. P. Bennett, P. Fouqué, A. Williams, M. Dominik, U. G. Jørgensen, D. Kubas, A. Cassan, C. Coutures, J. Greenhill, K. Hill, J. Menzies, P. D. Sackett, M. Albrow, S. Brillant, J. A. R. Caldwell, J. J. Calitz, K. H. Cook, E. Corrales, M. Desort, S. Dieters, D. Dominis, J. Donatowicz, M. Hoffman, S. Kane, J.-B. Marquette, R. Martin, P. Meintjes, K. Pollard, K. Sahu, C. Vinter, J. Wambsganss, K. Woller, K. Horne, I. Steele, D. M. Bramich, M. Burgdorf, C. Snodgrass, M. Bode, A. Udalski, M. K. Szymański, M. Kubiak, T. Więckowski, G. Pietrzyński, I. Soszyński, O. Szewczyk, Ł. Wyrzykowski, B. Paczyński, F. Abe, I. A. Bond, T. R. Britton, A. C. Gilmore, J. B. Hearnshaw, Y. Itow, K. Kamiya, P. M. Kilmartin, A. V. Korpela, K. Masuda, Y. Matsubara, M. Motomura, Y. Muraki, S. Nakamura, C. Okada, K. Ohnishi, N. J. Rattenbury, T. Sako, S. Sato, M. Sasaki, T. Sekiguchi, D. J. Sullivan, P. J. Tristram, P. C. M. Yock, and T. Yoshioka. Discovery of a cool planet of 5.5 Earth masses through gravitational microlensing. *Nature*, 439:437–440, January 2006. doi: 10.1038/nature04441.

F. W. Dyson, A. S. Eddington, and C. Davidson. A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919. *Royal Society of London Philosophical Transactions Series A*, 220:291–333, 1920.

A. S. Eddington. The total eclipse of 1919 May 29 and the influence of gravitation on light. *The Observatory*, 42:119–122, March 1919.

A. S. Eddington. On the instability of Einstein's spherical world. *Monthly Notices of the R.A.S.*, 90:668–678, May 1930.

A. Einstein. Zur Elektrodynamik bewegter Körper. *Annalen der Physik*, 322:891–921, 1905. doi: 10.1002/andp.19053221004.

A. Einstein. Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik*, 354:769–822, 1916. doi: 10.1002/andp.19163540702.

A. Einstein. Lens-Like Action of a Star by the Deviation of Light in the Gravitational Field. *Science*, 84:506–507, December 1936. doi: 10.1126/science.84.2188.506.

A. Friedman. Über die Krümmung des Raumes. *Zeitschrift fur Physik*, 10:377–386, December 1922. doi: 10.1007/BF01332580.

E. Hubble. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, 15:168–173, March 1929. doi: 10.1073/pnas.15.3.168.

G. Lemaître. Un Univers homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-galactiques. *Annales de la Societe Scietifique de Bruxelles*, 47:49–59, 1927.

J. Liesenborgs, S. de Rijcke, H. Dejonghe, and P. Bekaert. A generalization of the mass-sheet degeneracy producing ring-like artefacts in the lens mass distribution. *Monthly Notices of the R.A.S.*, 386:307–312, May 2008. doi: 10.1111/j.1365-2966.2008.13026.x.

J. A. Peacock, S. Cole, P. Norberg, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. D. Cannon, M. Colless, C. Collins, W. Couch, G. Dalton, K. Deeley, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, W. J. Percival, B. A. Peterson, I. Price, W. Sutherland, and K. Taylor. A measurement of the cosmological mass density from clustering in the 2dF Galaxy Redshift Survey. *Nature*, 410:169–173, March 2001.

S. Refsdal. The gravitational lens effect. *Monthly Notices of the R.A.S.*, 128:295–+, 1964a.

S. Refsdal. On the possibility of determining Hubble's parameter and the masses of galaxies from the gravitational lens effect. *Monthly Notices of the R.A.S.*, 128: 307–+, 1964b.

S. Refsdal. On the possibility of determining the distances and masses of stars from the gravitational lens effect. *Monthly Notices of the R.A.S.*, 134:315–+, 1966.

H. P. Robertson. Kinematics and World-Structure. *APJ*, 82:284–+, November 1935. doi: 10.1086/143681.

H. P. Robertson. Kinematics and World-Structure II. *APJ*, 83:187–+, April 1936a. doi: 10.1086/143716.

H. P. Robertson. Kinematics and World-Structure III. *APJ*, 83:257–+, May 1936b. doi: 10.1086/143726.

P. Saha. Lensing Degeneracies Revisited. *Astrophysical Journal*, 120:1654–1659, October 2000. doi: 10.1086/301581.

P. Saha and L. L. R. Williams. A Portable Modeler of Lensed Quasars. *Astrophysical Journal*, 127:2604–2616, May 2004. doi: 10.1086/383544.

P. Saha and L. L. R. Williams. Gravitational Lensing Model Degeneracies: Is Steepness All-Important? *APJ*, 653:936–941, December 2006. doi: 10.1086/508798.

A. G. Walker. On the formal comparison of Milne's kinematical system with the systems of general relativity. *Monthly Notices of the R.A.S.*, 95:263–269, January 1935.

R. J. Weymann, F. H. Chaffee, Jr., N. P. Carleton, D. Walsh, R. F. Carswell, and M. Davis. Multiple-mirror telescope observations of the twin QSOs 0957+561 A, B. *APJ Letters*, 233:L43–L46, October 1979. doi: 10.1086/183073.

2 Paper I

The Hubble Time Inferred from 10 Time Delay Lenses

...whereby we show the Universe is old enough to take care of itself.

The Astrophysical Journal, 650: L17–L20, 2006 October 10

THE HUBBLE TIME INFERRED FROM 10 TIME DELAY LENSES

Prasenjit Saha,[1,2] Jonathan Coles,[1] Andrea V. Macciò,[1] and Liliya L. R. Williams[3]

Received 2006 May 18; accepted 2006 July 11; published 2006 August 29

ABSTRACT

We present a simultaneous analysis of 10 galaxy lenses having time delay measurements. For each lens, we derive a detailed free-form mass map, with uncertainties, and with the additional requirement of a shared value of the Hubble parameter across all the lenses. We test the prior involved in the lens reconstruction against a galaxy-formation simulation. Assuming a concordance cosmology, we obtain $H_0^{-1} = 13.5^{+2.5}_{-1.3}$ Gyr.

Subject headings: cosmological parameters — galaxies: general — gravitational lensing

1. INTRODUCTION

If an object at cosmological distance is lensed into multiple images, the light-travel time for individual images differs. For variable sources, the differences are observable as time delays. The delays are of order

$$\Delta t \sim \frac{GM}{c^3} \sim (\Delta\theta)^2 H_0^{-1}, \tag{1}$$

where M is the lens mass and $\Delta\theta$ is the image separation (in radians). As Refsdal (1964) first pointed out, the effect provides an independent way of measuring H_0^{-1}. Time delay measurements have made much progress over the past decade, and now at least 15 are available (details below).

While equation (1) provides the order of magnitude, to determine the precise factor relating time delays and H_0^{-1}, one has to model the mass distribution. An observed set of image positions, rings, magnification ratios, and time delays is generically reproducible by many different mass models. This results in a large model-dependent uncertainty on the inferred Hubble parameter, even with perfect lensing data. To appreciate how serious this model dependence is, compare the models of B0957+561 by Kundić et al. (1997) and Bernstein & Fischer (1999): the results are $H_0 = 64 \pm 13$ and 77^{+29}_{-24} km s^{-1} Mpc^{-1}, respectively, both at 95% confidence; the more general models in the latter paper yield *larger* error bars. Alternatively, consider the nice summary in Figure 12 of Courbin (2006) of published H_0 estimates and uncertainties from individual lenses. Among the lenses shown, B1608+656 has all three of its independent time delays measured, B1115+080 has two delays measured, whereas the others have one each. One would expect these two best-measured lenses to be the best constrained. Yet B1608+656 has the largest error bars on H_0 and B1115+080 the second largest. This suggests that in the less-constrained lenses the real uncertainties are much larger but have been underestimated because the fewness of constraints did not force sufficient exploration of model dependence.

A general strategy for dealing with the nonuniqueness problem is to search through a large ensemble of models that can all reproduce the observations (Williams & Saha 2000; Oguri et al. 2004; Jakobsson et al. 2005). In this Letter, we follow such a strategy, simultaneously modeling 10 time delay lenses coupled by a shared Hubble parameter. The basic method is the same as in Saha & Williams (2004) and the accompanying PixeLens code, but a number of refinements have been made.

2. MODELING THE LENSES

Table 1 summarizes the lenses we have used. By “type” we mean the image morphology (Saha & Williams 2003): AD = axial double, ID = inclined double, SQ = short-axis quad, LQ = long-axis quad, and IQ = inclined quad. In B0957+561, two distinct source elements can be identified; both are lensed into ID.

We use PixeLens to generate an ensemble of 200 models. Each model in the ensemble consists of 10 pixelated mass maps and a shared value of H_0^{-1}. In addition to reproducing all the observed image positions and time delays, the mass maps are required to satisfy a prior. Errors in the image positions and time delays are assumed negligible, since they are much smaller than the range of models that reproduce the data. The details and justification of the prior are given in § 2 of Saha & Williams (2004), but basically the mass maps have to be nonnegative and centrally concentrated with a radial profile steeper than $|\theta|^{-0.5}$, since the lenses are galaxies. With one exception, the mass maps are required to have 180° rotation symmetry; only B1608+656 is allowed to be asymmetric, because the lens is known to contain two galaxies. A constant external shear is allowed for the lenses where the morphology shows evidence of external shear (all except B1608+656, B1104−181, and B2149−274). The lensing galaxies in B0957+561 and J0911+055 have cluster environments, but we have not treated these lenses differently. A concordance cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ is assumed.

We have not included magnification ratios as a constraint, for two reasons: first, optical flux ratios may be contaminated by microlensing (Keeton et al. 2006) and differential extinction; second, even tensor magnifications—that is, relative magnifications along different directions inferred from VLBI maps—are very weakly coupled with time delays (Raychaudhury et al. 2003), because magnification measures the local second derivative of the arrival time. Stellar velocity dispersions are available for some of the lenses, but we do not attempt to incorporate them, because current methods for doing so depend on strong assumptions about the mass distribution (Koopmans et al. 2006).

There are five additional candidates we have postponed modeling. B1830−211 has a time delay measurement (Lovell et al. 1998), but the lens position is uncertain (Courbin et al. 2002; Winn et al. 2002). B0909+532 (Ullán et al. 2006) also has an uncertain galaxy position. For B0435−122 (Kochanek et al. 2006) and J1131−123 (Morgan et al. 2006), our preliminary modeling appeared to imply asymmetric lenses, whereas the

---

[1] Institute for Theoretical Physics, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland.

[2] Astronomy Unit, Queen Mary and Westfield College, University of London, London E1 4NS, UK.

[3] Department of Astronomy, University of Minnesota, 116 Church Street, SE, Minneapolis, MN 55455.

TABLE 1

Lenses and Time Delays

Object	Type	z_L	z_S	Δt (days)
J0911+055	SQ	0.77	2.80	146 ± 8[a]
B1608+656	IQ	0.63	1.39	32 ± 2, 5 ± 2, 40 ± 2[b]
B1115+080	IQ	0.31	1.72	13 ± 2, 11 ± 2[c,d]
B0957+561	2 × ID	0.36	1.41	423 ± 1[e]
B1104−181	AD	0.73	2.32	161 ± 7[f]
B1520+530	ID	0.71	1.86	130 ± 3[g]
B2149−274	AD	0.49	2.03	103 ± 12[h]
B1600+434	ID	0.42	1.59	51 ± 4[i]
J0951+263	ID	0.24[j]	1.24	16 ± 2[j]
B0218+357	ID	0.68	0.96	10 ± 1[k,l]

[a] From Hjorth et al. (2002).
[b] From Fassnacht et al. (2002).
[c] From Schechter et al. (1997).
[d] From Barkana (1997).
[e] From Oscoz et al. (2001).
[f] From Ofek & Maoz (2003).
[g] From Burud et al. (2002b).
[h] From Burud et al. (2002a).
[i] From Burud et al. (2000).
[j] From Jakobsson et al. (2005).
[k] From Biggs et al. (1999).
[l] From Cohen et al. (2000).

image morphologies suggest fairly symmetric lenses. Finally, J1650+425 had its time delay measured (Vuissoz et al. 2006) as this Letter was being peer-reviewed.

We remark that while PixeLens in scientific terms is essentially the same as in Saha & Williams (2004), it has undergone several technical improvements. The key parameter in the code's performance is the total number of pixels (not pixels per lens), say, P. The memory required scales as P^2 and the time scales as P^3. The maximum usable P is in practice dictated not by time or memory but by the accumulation of round-off error. Our earlier paper attempted only three or four lenses at a time, going up to $P \simeq 600$. After improving the control of round-off error, PixeLens can now go up to $P \simeq 2000$ and beyond without difficulty. Meanwhile improving the memory management and implementing multithreading (which parallelizes the computation if run on a shared-memory multiprocessor machine) and newer hardware have more than compensated for the P^3 increase in arithmetic.

We have previously done two different tests of the general method. In Saha et al. (2006), the algorithm is tested by feeding time delays sampled from a model ensemble back into PixeLens and then recovering the model H_0^{-1}. This showed that any biases introduced by the ensemble-generating process have affected H_0^{-1} by less than 5% but did not test the prior. Williams & Saha (2000) presented a blind test where one author simulated data using simple model galaxies and a secret fictional value of H_0, and the other author recovered that value within uncertainties using an ancestor of PixeLens. That provided a basic test of the whole procedure, including the prior, but still assumed that the models chosen by the first author for the test were representative of real lensing galaxies. A similar test using current galaxy-formation simulations is desirable but technically formidable; however, we carry out a simple version of such a test below.

3. RESULTS

Our H_0^{-1} distribution is shown in Figure 1 and may be summarized as

$$H_0^{-1} = 13.5^{+2.5}_{-1.2}\ \text{Gyr} \quad (H_0 = 72^{+8}_{-11}\ \text{km s}^{-1}\ \text{Mpc}^{-1}) \qquad (2)$$

at 68% confidence and $13.5^{+5.6}_{-1.6}$ Gyr at 90% confidence. This estimate neglects measurement errors in the time delays. However, we have verified by repeating the analysis with perturbed time delays that the effect of measurement errors is very small. Astrometric errors are also very small.

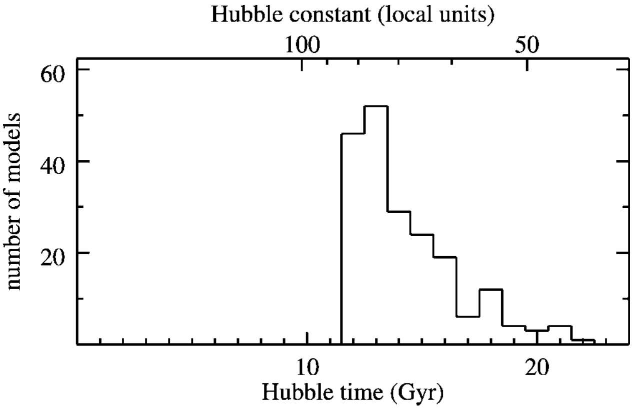


Fig. 1.—Histogram of the ensemble of H_0^{-1} values. The unbinned distribution gives $H_0^{-1} = 13.5^{+2.5}_{-1.2}$ Gyr at 68% confidence and $13.5^{+5.6}_{-1.6}$ Gyr at 90% confidence.

Figure 1 is consistent with the analogous Figures 8 and 11 in Saha & Williams (2004), which derive from two time delay quads and four doubles considered separately. But the constraints do not improve as much as simple $1/\sqrt{N}$ would predict. In fact, the uncertainties are far from Gaussian, and some lensing configurations are much more useful than others. Saha et al. (2006) discuss this point in more detail and conclude that a 5% uncertainty on H_0^{-1} is possible using 11 lenses, provided the lenses all have favorable configurations.

Figure 2 shows ensemble average mass distributions for the 10 lenses. Notice that some lenses, especially B1115+080, B1104−181, and B1520+530, have twisting isodensity contours and/or radially dependent ellipticities, features that are not included in parameterized models.

The lens galaxies have varying amounts of dark matter. This follows from Ferreras et al. (2005), who compare the total-mass profiles of 18 lensing galaxies, including six from the present sample, with stellar-mass profiles from population-evolution models. (The work assumed $H_0^{-1} = 14$ Gyr, which is well within our uncertainties, and hence the results are valid for the models here.) From their Table 1, we see that out to ∼$3R_{\text{eff}}$, B1520+530 is mainly stars, B1115+080, B1608+656, and B2149−274 have significant nonstellar mass, while J0951+263 and B1104−181 are dominated by dark matter.

4. LENS MODELS COMPARED WITH A SIMULATION

We now address a simplified version of the question: are our lens models typical of current galaxy-formation simulations?

The details of gas dynamics, star formation, active galactic nucleus formation, and feedback on galaxy scales are still uncertain. With this caveat in mind, we consider a single high-resolution galaxy, extracted from an N-body cosmological simulation and then resimulated using the TreeSPH GASOLINE code (Wadsley et al. 2004) including gas cooling, star formation, supernova feedback, and UV background. The galaxy is an E1 or E2 triaxial elliptical dominated by stars in the inner region but overall ∼80% dark matter (Macciò et al. 2006). Orienting this galaxy randomly and ray-tracing with random sources (Macciò 2005), we generated about 500 quads and 10,000 doubles, and calculated time delays for each of these.

As equation (1) suggests, time delays generated from a single

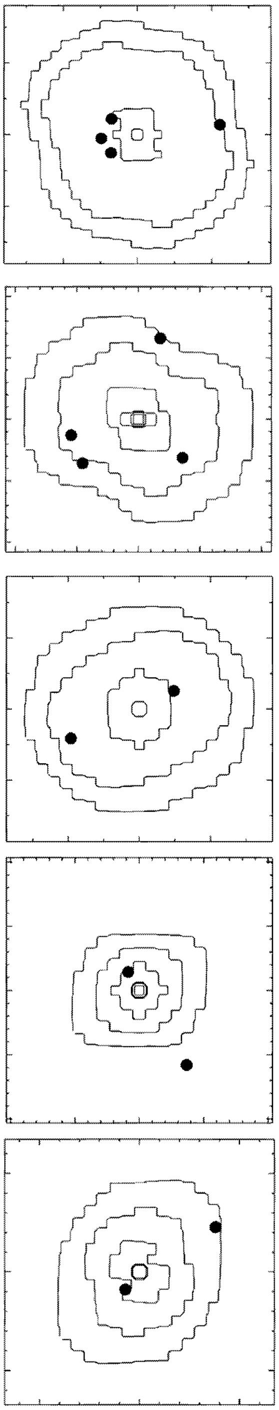


FIG. 2.—Ensemble-average mass maps of the lenses: J0911+055 (*upper left*), B1608+656 (*upper right*), B1115+080, B0957+561, B1104−181, B1520+530, B2149−274, B1600+434, J0951+263, and B0218+357. The larger tick marks in each panel correspond to 1″. The contours are in logarithmic steps, with critical density corresponding to the third contour from the outside.

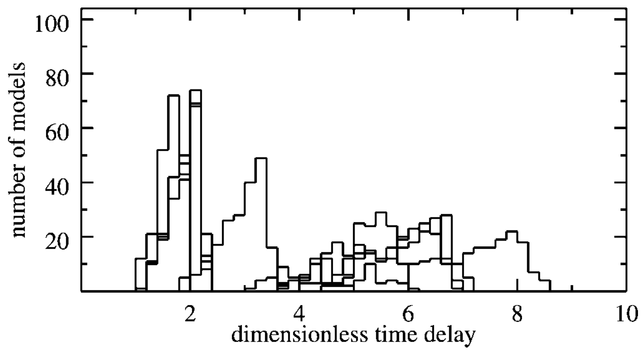


FIG. 3.—Histograms of φ for all 10 lenses. J0911+055, B1608+656, and B1115+080 all peak around 2. B0957+561 peaks around 3. B1104−181 peaks around 2. B1520+530, B1600+434, J0951+263, and B0218+357 all peak around 6. B2149−274 peaks around 8.

galaxy will range over a factor of only a few and cannot be directly compared with the observed time delays, which range over a factor of 40. We therefore consider a dimensionless form of the time delay φ, given by

$$H_0 \Delta t = \varphi \frac{1}{16} (\theta_1 + \theta_2)^2 D, \qquad (3)$$

where θ_1, θ_2 are the lens-centric distances (in radians) of the first and last images to arrive, Δt is the observed time delay between them, and D is the usual distance factor in lensing. This factors out the dependence of the time delay on cosmology (through H_0 and D) and on the scale of the lens [through $(\theta_1 + \theta_2)^2$], leaving φ dependent only on the shape and steepness of the lens and on the source position with respect to the caustics (Saha 2004).

Figure 3 shows the histograms of φ in our lens-model ensembles. The quads all peak around 2, while the doubles mostly peak around 5–8; the exceptions are B0957+561 peaking around 3 and B1104−181 peaking around 2. Since B0957+561 is in a cluster, it is plausible that the mass profile is unusually shallow, thus reducing the time delay through the well-known steepness degeneracy. The low value for B1104−181 is more puzzling. Figure 4 is simpler, showing the probability distributions of φ for doubles and quads generated by the single simulated galaxy.

Figures 3 and 4 are not quite equivalent, but we can think of both as derived from an underlying prob(φ | galaxy, lensing obs). Each histogram in Figure 3 weights this probability distribution by observation selection effects and by the PixeLens prior, and then marginalizes over galaxies while holding the lensing observables fixed. Figure 4 marginalizes over lensing observables (separately for doubles and quads) while holding the galaxy fixed. Bearing this difference in mind, the simulated galaxy appears typical of our lens models. The most noticeable difference is the absence of observed quads with φ close to zero; but that is an expected observational selection effect, because very short time delays are unlikely to be measured.

We conclude that the PixeLens prior, as far as this preliminary experiment can reveal, is consistent with galaxy-formation sim-

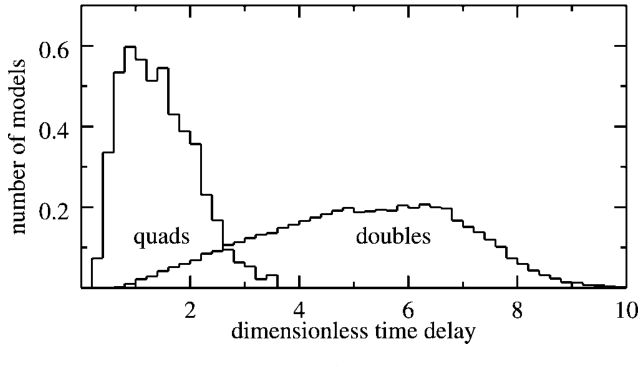


FIG. 4.—Probability distribution of φ for the simulated galaxy. Doubles and quads are normalized separately.

ulations. Further comparisons with simulated galaxies and fine-tuning of the prior are desirable in future work.

5. DISCUSSION

We have expressed our main result (Fig. 1) preferentially in terms of H_0^{-1} rather than H_0 because the former appears more naturally in lensing theory. But it is interesting to continue with H_0^{-1} in comparing with other techniques, because H_0^{-1} has a simple interpretation quite generally: it is $a/\dot{a}$ or the doubling time for metric distances at the current expansion rate. Coincidentally,[4] in the concordance cosmology ($K = 0$, $\Omega_m \simeq \frac{1}{4}$, $w = -1$) H_0^{-1} also equals the expansion age of the universe, within uncertainties. In particular, H_0^{-1} estimates can be immediately compared with globular cluster ages, such as in Krauss & Chaboyer (2003).

The well-known recent measurements of H_0^{-1}, expressed in gigayears, are:

1. 13.6 ± 1.5 from Freedman et al. (2001), who combine several different indicators calibrated using Cepheids;
2. 15.7 ± 0.3 (statistical) ± 1.2 (systematic) from Sandage et al. (2006), who use Type Ia supernova distances calibrated using Cepheids;
3. 13.6 ± 0.6 from Spergel et al. (2006), who use the cosmic microwave background (CMB) fluctuation spectrum.

Our result is consistent with any of these.

It is worth noting, however, that the Hubble parameter appears in very different guises in different techniques. The distance-ladder methods measure the local cosmological expansion rate, independent of the global geometry. By contrast, in the CMB, H_0^{-1} is one parameter in a global cosmological model. Lensing is different again: here one assumes a global geometry and then measures a single scale parameter. The same is true of Sunyaev-Zel'dovich and X-ray clusters. The latter technique has made significant progress recently (Jones et al. 2005) but thus far still relies on strong assumptions: spherical symmetry of the cluster potential and hydrostatic equilibrium of the gas. In principle, lensing time delays can determine the global geometry as well (Refsdal 1966), but the amount of data needed is not observationally viable yet.

Whether lensing time delays can get the uncertainties in the Hubble parameter down to the 5% level is an open question. Maybe galaxy-lens models can be constrained enough to determine H_0^{-1} to better than 5%, thus making lensing the preferred method (Schechter 2004); or maybe the approach is best used in reverse, inputting H_0^{-1} to constrain galaxy structure (Kochanek et al. 2006). Fortunately, either outcome is worthwhile, and the basic technique will be the same. So whether the optimists or the pessimists are right, the usual cliches of "more data!" (time delay measurements) and "more theory!" (lens models) are both apt.

---

[4] Although a spoof paper by Scott (2006) develops a conspiracy theory for this.

REFERENCES

Barkana, R. 1997, ApJ, 489, 21
Bernstein, G., & Fischer, P. 1999, AJ, 118, 14
Biggs, A. D., Browne, I. W. A., Helbig, P., Koopmans, L. V. E., Wilkinson, P. N., & Perley, R. A. 1999, MNRAS, 304, 349
Burud, I., et al. 2000, ApJ, 544, 117
———. 2002a, A&A, 383, 71
———. 2002b, A&A, 391, 481
Cohen, A. S., Hewitt, J. N., Moore, C. B., & Haarsma, D. B. 2000, ApJ, 545, 578
Courbin, F. 2006, in ASP Conf. Ser. 326, Gravitational Lensing: A Unique Tool for Cosmology, ed. D. Valls-Gabaud & J.-P. Kneib (San Francisco: ASP), in press (astro-ph/0304497)
Courbin, F., Meylan, G., Kneib, J.-P., & Lidman, C. 2002, ApJ, 575, 95
Fassnacht, C. D., Xanthopoulos, E., Koopmans, L. V. E., & Rusin, D. 2002, ApJ, 581, 823
Ferreras, I., Saha, P., & Williams, L. L. R. 2005, ApJ, 623, L5
Freedman, W. L., et al. 2001, ApJ, 553, 47
Hjorth, J., et al. 2002, ApJ, 572, L11
Jakobsson, P., Hjorth, J., Burud, I., Letawe, G., Lidman, C., & Courbin, F. 2005, A&A, 431, 103
Jones, M. E., et al. 2005, MNRAS, 357, 518
Keeton, C. R., Burles, S., Schechter, P. L., & Wambsganns, J. 2006, ApJ, 639, 1
Kochanek, C. S., Morgan, N. D., Falco, E. E., McLeod, B. A., Winn, J. N., Dembicky, J., & Ketzeback, B. 2006, ApJ, 640, 47
Koopmans, L. V. E., Treu, T., Bolton, A. S., Burles, S., & Moustakas, L. A. 2006, ApJ, in press (astro-ph/0601628)
Krauss, L. M., & Chaboyer, B. 2003, Science, 299, 65
Kundić, T., et al. 1997, ApJ, 482, 75
Lovell, J. E. J., Jauncey, D. L., Reynolds, J. E., Wieringa, M. H., King, E. A., Tzioumis, A. K., McCulloch, P. M., & Edwards, P. G. 1998, ApJ, 508, L51
Macciò, A. V. 2005, MNRAS, 361, 1250
Macciò, A. V., Moore, B., Stadel, J., & Diemand, J. 2006, MNRAS, 366, 1529
Morgan, N. D., Kochanek, C. S., Falco, E. E., & Dai, X. 2006, ApJ, submitted (astro-ph/0605321)
Ofek, E. O., & Maoz, D. 2003, ApJ, 594, 101
Oguri, M., et al. 2004, ApJ, 605, 78
Oscoz, A., et al. 2001, ApJ, 552, 81
Raychaudhury, S., Saha, P., & Williams, L. L. R. 2003, AJ, 126, 29
Refsdal, S. 1964, MNRAS, 128, 307
———. 1966, MNRAS, 132, 101
Saha, P. 2004, A&A, 414, 425
Saha, P., Courbin, F., Sluse, D., Dye, S., & Meylan, G. 2006, A&A, 450, 461
Saha, P., & Williams, L. L. R. 2003, AJ, 125, 2769
———. 2004, AJ, 127, 2604
Sandage, A., Tammann, G. A., Saha, A., Reindl, B., Macchetto, F. D., & Panagia, N. 2006, ApJ, submitted (astro-ph/0603647)
Scott, D. 2006, preprint (astro-ph/0604011)
Schechter, P. L. 2004, in IAU Symp. 225, The Impact of Gravitational Lensing on Cosmology, ed. Y. Mellier & G. Meylan (Cambridge: Cambridge Univ. Press), 281
Schechter, P. L., et al. 1997, ApJ, 475, L85
Spergel, D. N., et al. 2006, ApJ, submitted (astro-ph/0603449)
Ullán, A., Goicoechea, L. J., Zheleznyak, A. P., Koptelova, E., Bruevich, V. V., Akhunov, T., & Burkhonov, O. 2006, A&A, 452, 25
Vuissoz, C., et al. 2006, A&A, submitted (astro-ph/0606317)
Wadsley, J. W., Stadel, J., & Quinn, T. 2004, NewA, 9, 137
Williams, L. L. R., & Saha, P. 2000, AJ, 119, 439
Winn, J. N., Kochanek, C. S., McLeod, B. A., Falco, E. E., Impey, C. D., & Rix, H.-W. 2002, ApJ, 575, 103

3 Paper II

A New Estimate of the Hubble Time with Improved Modeling of Gravitational Lenses

...whereby we show the Universe is still old enough to take care of itself,
and can also make a nice cup of tea.

The Astrophysical Journal, 679:17–24, 2008 May 20

A NEW ESTIMATE OF THE HUBBLE TIME WITH IMPROVED MODELING OF GRAVITATIONAL LENSES

Jonathan Coles
Institut für Theoretische Physik, Universität Zürich, Winterthurerstrasse 190, CH-8057, Zürich, Switzerland; jonathan@physik.uzh.ch
Received 2007 June 25; accepted 2008 February 18

ABSTRACT

This paper examines free-form modeling of gravitational lenses using Bayesian ensembles of pixelated mass maps. The priors and algorithms from previous work are clarified and significant technical improvements are made. Lens reconstruction and Hubble time recovery are tested using mock data from simple analytic models and recent galaxy formation simulations. Finally, using published data, the Hubble time is inferred through the simultaneous reconstruction of 11 time delay lenses. The result is $H_0^{-1} = 13.7^{+1.8}_{-1.0}$ Gyr ($H_0 = 71^{+6}_{-8}$ km s^{-1} Mpc^{-1}).

Subject headings: cosmological parameters — gravitational lensing

1. INTRODUCTION

Gravitational lenses (GLs) provide a fantastic natural tool for probing many of the large-scale properties of the cosmos. Recent applications range from estimating the age of the universe (Saha et al. 2006) to studying the dark matter profiles of galaxies (Read et al. 2007) to testing alternative theories of gravity (Ferreras et al. 2007).

Despite their potential, GLs are difficult to study because of several degeneracies such as the position of the source and the mass distribution of the lensing object. This paper focuses on strong lensing of quasars by galaxies, but the techniques developed can equally be applied to clusters. Many have tried to fit models to GLs by assuming different galaxy structures. Young et al. (1981) were the first to do so with King models, and many others have followed using a variety of single isothermal ellipses, Sèrsic models, or de Vaucouleurs profiles (for a review, see Kochanek 2004). But different models can easily give different results (Vuissoz et al. 2007).

This kind of modeling is generally called parametric modeling. Each model has a nominal amount of parameters that can be adjusted. But while one model may fit the data, the degeneracies make it difficult to determine how well these models really represent the lens; and as pointed out by Bernstein & Fischer (1999) and more recently by Read et al. (2007) in connection with time delays, without extreme care these models can be very sensitive to the assumptions.

In contrast, free-form or nonparametric models reconstruct the lens on a grid or a set of basis functions. No particular form is assumed, and the results allow a wider range of solutions than parametric models might. Such modeling is not unique to lensing, though.

Schwarzschild (1979) used nonparametric modeling to show for the first time that it is possible to construct a triaxial stellar system in equilibrium. He showed that there existed a distribution of stars on orbits that fit a given density function D. The three-dimensional space of a galaxy was divided into M cells, and D was expressed by $D(J) = \sum_{I=1}^{M} C(I)B(I,J)$, where $B(I,J)$ is the orbit density for an orbit I in cell J, calculated using test particles in a fixed potential, and $C(I)$, the number of stars on orbit I, was determined numerically by solving a linear program.

In a very similar manner, Schwarzschild's technique can be applied to lenses. Modeling the lenses on a grid was first introduced by Saha & Williams (1997) and then later extended to include both weak and strong lensing by AbdelSalam et al. (1998). Similar methods have also been used by Diego et al. (2005) and Bradač et al. (2005). But in contrast to Schwarzschild, it is desirable to show the *variety* of solutions rather than just existence. This important feature is incorporated into the work of Williams & Saha (2000) and the software PixeLens (Saha & Williams 2004). Related approaches are developed in Trotter et al. (2000) and Keeton & Winn (2003). Given a large ensemble of models, one or several variables are examined while averaging out (marginalizing) the others. The same principle is used in statistical mechanics. However, the use of marginalization is sometimes overlooked, leading to a misunderstanding that pixelated models are "grossly underconstrained" because the number of variables exceeds the number of data points.

Pixelated modeling has the advantage of allowing the form of the lens to vary. It does not presuppose important parameters and can produce models that would otherwise not be possible with parametric modeling. For instance, while parametric models already showed that steepness is an important parameter (Wambsganss & Paczynski 1994), pixelated models showed that shape degeneracies, which are often difficult to capture with parametric models, cannot be ignored (Saha & Williams 2006); twists and nonuniform stretching are also easily found.

In this paper, pixelated lens modeling and the constraints imposed on the models are explicitly defined. The algorithms are improved with several optimization techniques and the enhanced method is tested against lenses from an N-body simulation and another fictitious data set. Finally, a system of 11 lenses is used in the same way as Saha et al. (2006) to further constrain the Hubble time.

2. CREATING MODELS

PixeLens generates an ensemble of lens models that fit the input data. In the Bayesian way, the ensemble itself provides estimates and uncertainties. Each model consists of a set of n discrete mass squares with density κ_n, a source position $\boldsymbol{\beta}$, and optionally, a variable h which is proportional to H_0. If the time delays are unknown, the value of h is fixed. In this paper, where the time delays are known, h varies across the ensemble. The positions of observed images and the redshifts of the source and lens are taken to be given with errors small enough to be ignored. Time delays between images, when available, are similarly assumed to be accurate. Tests from Saha et al. (2006) show that adjusting these numbers slightly to simulate errors has much less effect than the model uncertainties.

The mass density in each square, or *pixel*, is the projected mass density on the plane of the sky in units of the critical density.[1] The pixelated surface is a disk of radius `pixrad` pixels. The total number of pixels is then about $\pi \times \mathtt{pixrad}^2$. The extent of the modeled mass, `maprad`, defaults to $\min(r_{\max} + r_{\min}, 2r_{\max} - r_{\min})$, where $r_{\min}$ and $r_{\max}$ are the distances of the innermost and outermost images, respectively. This allows for a buffer zone outside the outermost image when required.

Following Blandford & Narayan (1986), the arrival time is the light travel time scaled by

$$h^{-1}T(z_L, z_S) = (1 + z_L)\frac{D_L D_S}{cD_{LS}}, \tag{1}$$

where z_L is the redshift of the lens, and D_L, D_S, and D_{LS} are the distances from observer to lens, observer to source, and lens to source, respectively. This removes the dependence on a particular cosmology. The h^{-1} factor comes through the distance factors.

The arrival time at position $\boldsymbol{\theta}$ is given by

$$\tau(\boldsymbol{\theta}) = \frac{1}{2}|\boldsymbol{\theta}|^2 - \boldsymbol{\theta} \cdot \boldsymbol{\beta} - \int \ln|\boldsymbol{\theta} - \boldsymbol{\theta}'|\kappa(\boldsymbol{\theta}')d^2\boldsymbol{\theta}'. \tag{2}$$

This can be interpreted as a surface, which is modeled with a summation over the pixel densities,

$$\tau(\boldsymbol{\theta}) = \frac{1}{2}|\boldsymbol{\theta}|^2 - \boldsymbol{\theta} \cdot \boldsymbol{\beta} - \sum_n \kappa_n Q_n(\boldsymbol{\theta}) + \gamma_1\left(\theta_x^2 - \theta_y^2\right) + 2\gamma_2\theta_x\theta_y. \tag{3}$$

Two additional terms involving γ_1 and γ_2 are added to account for external shear from neighboring galaxies.

The function Q is the integral from equation (2) evaluated over a square pixel with side length a and is defined using the same notation as in Saha & Williams (1997). Let x, y be the Cartesian components of $\boldsymbol{\theta}$, $r^2 = x^2 + y^2$, and

$$\begin{aligned}\tilde{Q}_n(x, y) = (2\pi)^{-1}\big[&x^2 \arctan(y/x) \\ &+ y^2 \arctan(x/y) + xy(\ln r^2 - 3)\big].\end{aligned} \tag{4}$$

Then

$$Q_n(x, y) = \tilde{Q}_n(x_+, y_+) + \tilde{Q}_n(x_-, y_-) - \tilde{Q}_n(x_-, y_+) - \tilde{Q}_n(x_+, y_-), \tag{5}$$

where $x_\pm = x - x_n \pm a/2$ and $y_\pm = y - y_n \pm a/2$.

The function τ is linear in all the unknowns $\boldsymbol{\beta}$, κ_n, γ_1, and γ_2. Constraints are placed on τ and the unknowns so that the results are physical. The data constraints come directly from lensing theory. The priors, or assumptions, are additional constraints that are physically motivated.

As a side note, the source position can be negative because it is relative to the center, but it must be positive in order to encode it as part of the linear program. This is resolved by adding a constant internally.

Data Constraint 1. *Images are observed where the arrival time surface is stationary,* $\nabla\tau(\boldsymbol{\theta}_i) = 0$ *(Fermat's principle),*

$$\begin{aligned}\boldsymbol{\theta}_{i,x} - \boldsymbol{\beta}_x - \sum dQ/d\boldsymbol{\theta}_{i,x} &= 0, \\ \boldsymbol{\theta}_{i,y} - \boldsymbol{\beta}_y - \sum dQ/d\boldsymbol{\theta}_{i,y} &= 0.\end{aligned} \tag{6}$$

Data Constraint 2. *The time delay between two images* $\boldsymbol{\theta}_i$ *and* $\boldsymbol{\theta}_j$ *must be consistent with observations,*

$$\tau(\boldsymbol{\theta}_i) - \tau(\boldsymbol{\theta}_j) = h\frac{(\text{obs delay})}{T(z_L, z_S)}. \tag{7}$$

If the time delays are unknown, the time ordering can be inferred from the morphology and imposed by

$$\tau(\boldsymbol{\theta}_i) - \tau(\boldsymbol{\theta}_j) \geq 0. \tag{8}$$

Data Constraint 3. *At each* $\boldsymbol{\theta}_i$ *there are two constraints of the form*

$$\epsilon\left|\frac{\partial^2}{\partial\theta_{x'}^2}\tau(\boldsymbol{\theta}_i)\right| \leq \left|\frac{\partial^2}{\partial\theta_{y'}^2}\tau(\boldsymbol{\theta}_i)\right|, \tag{9}$$

where $\theta_{x'}$ and $\theta_{y'}$ are the local radial and tangential directions and $\epsilon = 1/10$ by default. This ensures that the image elongation is between ϵ and $1/\epsilon$ when projected along the radial direction. In practice, the default does not place any constraints on the image. If an image is known to be elongated, then ϵ can be changed. In particular, this was used in AbdelSalam et al. (1998).

Data Constraint 4. *If a model contains N lenses, they must share the same Hubble constant,*

$$h_{\text{lens}_1} = h_{\text{lens}_2} = \ldots = h_{\text{lens}_N}. \tag{10}$$

When H_0 is unspecified, then H_0 is allowed to vary from model to model but not from lens to lens within a single model.

The following priors are the assumptions made about the lensing systems. All are well defined and astrophysically justified, as explained below.

Prior 1. *The density cannot be negative,*

$$\kappa_n \geq 0. \tag{11}$$

This is a quite trivial requirement, but one that can often be difficult to ensure with other techniques. The linear programming algorithm employed here guarantees this prior by design.

Notice the similarity between Schwarzschild's equation from § 1 on the one hand and equation (3) and prior 2. There is a linear function (D or τ) whose value is known and a summation over a product where one of the product terms is calculated beforehand (B or Q). Schwarzschild was limited at the time to what he could say about the unknowns, but negative values were not allowed. The goal was only to show the existence of one solution, because no one knew at the time whether a triaxial solution was possible. With lenses, much more can be said about the unknowns and lensing is known to occur.

Prior 2. *Most lens are assumed to have inversion symmetry, unless the lenses are observed to be interacting or otherwise strongly asymmetric,*

$$\kappa_{i,j} = \kappa_{-i,-j}. \tag{12}$$

---

[1] Many have suggested that it would be better to discretize the potential, but the potential is not naturally discrete and doing so would require recovering the mass from Poisson's equation; guaranteeing that the mass remains positive is difficult and involves a double derivative which produces noisy results.

Prior 3. *The density gradient should point within* $\theta = 45^\circ$ *of the center*,

$$\begin{aligned}[\, i \quad j \,]\mathbf{M}\nabla\kappa_{i,j} &\geq 0, \\ [\, i \quad j \,]\mathbf{M}^T\nabla\kappa_{i,j} &\geq 0, \end{aligned} \tag{13}$$

where

$$\mathbf{M} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \tag{14}$$

$$\nabla\kappa_{i,j} \equiv (2a)^{-1}(\kappa_{i+1,j} - \kappa_{i-1,j} - \kappa_{i,j+1} - \kappa_{i,j-1}), \tag{15}$$

and a is the pixel size. This complicated expression is just saying that if the density gradient of a pixel were pointing at most θ away from the center then moving the pixel's position by θ should align the density gradient so that it points directly at the center. If the gradient is greater than θ the greater than or equal to condition will not be satisfied.

Prior 4. *The density of a pixel must be no more than twice the average density of its neighbors*,

$$\kappa_n \leq 2\frac{1}{N(n)}\sum_{i\in N(n)}\kappa_i, \quad n \neq 1. \tag{16}$$

This is a weak smoothing criterion. Normally, it is not applied to the central pixel, which can have arbitrary density.

Prior 5. *The mass profile must be steeper than* r^{-s}. *Let* R_i *be the set of all pixels on a discretized "ring" i of radius* r_{R_i}, *one pixel thick. The number of pixels in a ring is* $|R_i|$. *Let* $C_i = r_{R_i}^s/|R_i|$, *then*

$$C_i\sum_{n\in R_i}\kappa_n - C_{i+1}\sum_{n\in R_{i+1}}\kappa_n \geq 0. \tag{17}$$

The default radial mass profile constraint has $s = 0.5$. This is intentionally rather shallow, but as explained in Saha & Williams (2004), this is motivated by evidence showing that total density distribution in central regions of ellipticals is roughly isothermal, i.e., r^{-2}. Furthermore, the projected gas density in the Milky Way scales as $r^{-1.75}$ (Binney et al. 1991).

Again, the most important thing to realize from the constraints and the discretized lens equation is that *the constraints are all linear*. They can therefore be solved using any number of linear programming techniques. However, rather than find one solution, the space of all solutions is sampled to understand the distribution.

2.1. *Bayesian Monte Carlo Markov Chain Sampling*

The linear equations presented in § 2 constrain the solution space to a convex multidimensional polyhedron known as a simplex. The interior points of the simplex are solutions to the linear equations.

PixeLens samples the interior points using a Monte Carlo Markov Chain (MCMC) technique. The general technique is described in condensed matter texts (Binney et al. 1992) and Bayesian books (Saha 2003). Each solution is used to reconstruct the arrival time surface, mass density contours, H_0^{-1}, etc.

The sampling method, Algorithm S, relies on being able to find random vertices of the simplex. The current implementation uses the standard linear programming simplex algorithm (Dantzig 1963; Press et al. 1986; Cormen et al. 2001) to maximize a given *objective function* subject to the linear constraints that form the simplex. The maximum is guaranteed to be at a vertex. For the present purposes, the objective function is chosen randomly after each iteration of Algorithm S, thereby producing a new vertex each time.

Algorithm S. (*Sample interior points*).

1. Let γ_0 be a vertex on the simplex and $i = 0$ the index of the current iteration.
2. Let α_i be a new vertex.
3. Extend a line from α_i through γ_i until a constraint is reached. Select an interior point γ_{i+1} uniformly from the line.
4. If another model is desired, increment i and go back to step 2, otherwise stop.

Because the simplex is convex by construction of linear hyperplanes, Algorithm S is guaranteed to return models in the solution space.

In addition to the explicit priors of § 2, there is also a prior imposed by the sampling strategy itself. Although clearly well defined, the physical significance of this prior continues to be the subject of study. This is not a point to be lightly dismissed, since it influences the derived distribution of H_0. However, the strategy cannot be arbitrary and there are very strict requirements on the way the volume can be sampled, which are discussed below. Numerous tests, both in this paper and others, have shown that the weighting can be empirically justified. The key point is that many different models must be examined. Other modeling techniques tend to assume the correctness of the model that is fit to the data, rather than letting the data itself reveal the model. To quote Blandford & Kundic (1997), "We should still aggressively explore all other classes of models that can also fit the observations but yet which produce disjoint estimates for the time delay. The true uncertainty in the Hubble constant is given by the union of all of these models."

Algorithm S, in effect, puts a metric on the union. Previous PixeLens papers implied that the sampling of the simplex was uniform in volume, but this is not correct, nor is it desired. The space does not have a Euclidean metric, and while it is still unclear what metric the space *should* have, there are certain properties that an algorithm sampling the space *must* have.

1. The sampling strategy must be insensitive to changes in dimensionality of the space. In other words, increasing the number of variables (e.g., by increasing the pixel resolution, which subdivides pixels) should not change the predicted values of H_0. This is not true if the solution space is uniformly sampled. As an example of the problem, imagine a uniformly sampled right triangle where the legs meet at the origin. The density of points projected onto one axis will be greater toward the origin. In higher dimensions, when the points are projected onto the same axis, the density distribution will be skewed further toward the origin.
2. The sampling strategy must be insensitive to units. The variables that define the solution space do not all have the same units. Some are mass density, some are source positions, one is H_0, etc. By simply scaling any of these units, the space is stretched or compressed. This would affect a sampling strategy based on volume when the number of dimensions is greater than two. Whatever the sampling prior is, it *must* be insensitive to this.

Both of these serious problems are solved by Algorithm S. The first problem is solved because a point is chosen uniformly along the line in step 3 regardless of the number of dimensions. One can see from Figure 1 that the predicted value of H_0 remains quite steady even as the number of variables is increasing. The second problem is solved because the vertices of the space are used to guide the sampling strategy. How the vertices are chosen is completely independent of units. Any scaling would not affect the vertex selection procedure. Figure 2 shows a three-dimensional

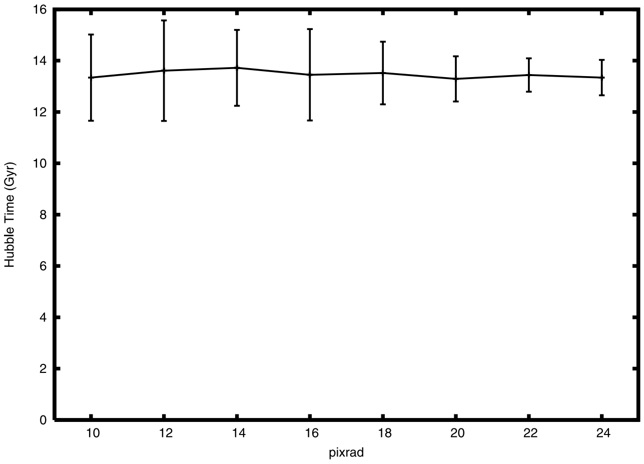


FIG. 1.—Predicted Hubble time as a function of the pixel radius of the grid. The number of variables is $O(\texttt{pixrad}^2)$. Error bars indicate the 1 σ deviations from the medians. Increasing the variables does not grossly affect the median H_0, showing that condition 1 of the sampling strategy is satisfied. A single lens, B1115+080, is being modeled.

sampled simplex. The sampling is clearly not uniform but is insensitive to stretching of the axes.

Algorithm S has changed slightly from older versions of PixeLens. Previous versions took a fixed number of vertex steps. The new vertex was often very close to the old one and resulted in clumps of correlated models. The new version seeks out vertices further away, which reduces the problem and better samples the interior with fewer samples. The running time increased with this change, but the results are more representative. Within the errors, however, old results are still valid.

Although Algorithm S does not sample the volume uniformly, in the limit of infinite samples, it does have *some* distribution. But how well is that distribution recovered with only a finite number of samples? To approximate the true distribution, 10,000 models of the lens B1115+080 were generated. The "finite" sample consisted of 200 models. Figure 3 compares the distribution of just the Hubble time variable. When the two samples are taken from the same distribution, the crosses fall on the dashed line. Even with a small sample, the distribution is well recovered.

2.2. *Technical Issues*

While PixeLens is stable, variations on sampling can introduce numerical instability. If a point is not chosen uniformly from the line in step 3 of Algorithm S, a numerical error in the coordinates of sampled points will grow exponentially fast and lead to future points lying *outside* the solution space. Reprojecting a point back into the space is impractical, because the exact size and shape of the simplex is unknown and truly incalculable due to the extraordinarily large number of dimensions and vertices. (It is worth noting that if all the vertices could be known in advance, the simplex algorithm would be unnecessary. One could simply pick a new vertex from the list.) In the worst case, however, this error is detectable. If such an error is detected, the program will issue a message and halt.

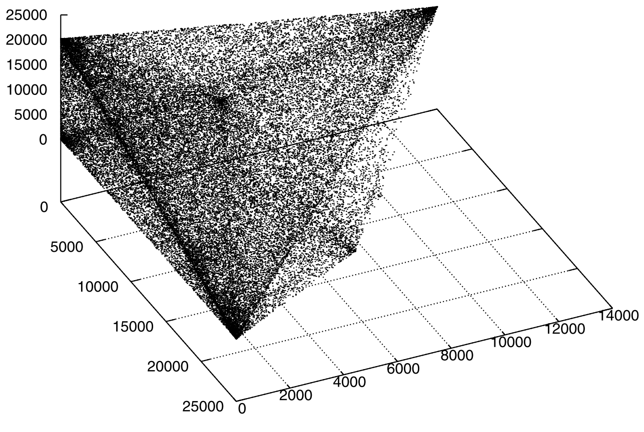


FIG. 2.—Three-dimensional example of a sampled simplex with 50,000 points. The overdensities clearly indicate that the volume is not uniformly sampled. This must be the case in order to satisfy the two conditions of the sampling strategy.

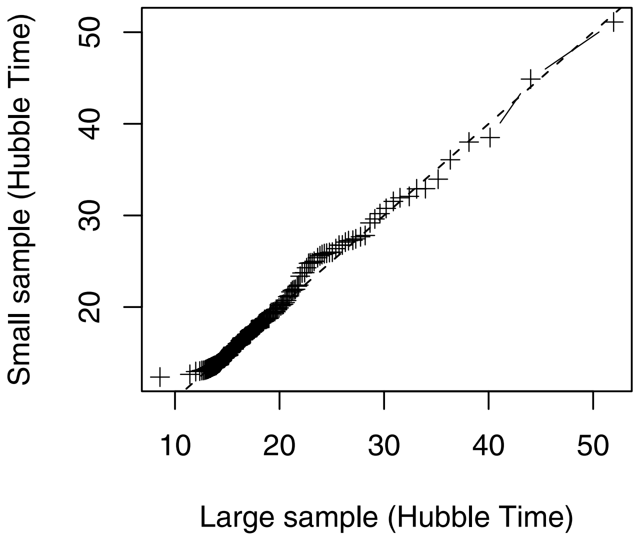


FIG. 3.—Quantile-quantile plot comparing the distribution of a large sample of Hubble times to the distribution of a small sample. The points lie nearly perfectly on the dashed line, indicating that the two samples come from the same distribution. The tail extends off the figure because of a few extreme outliers in the large sample. The figure was clipped for clarity.

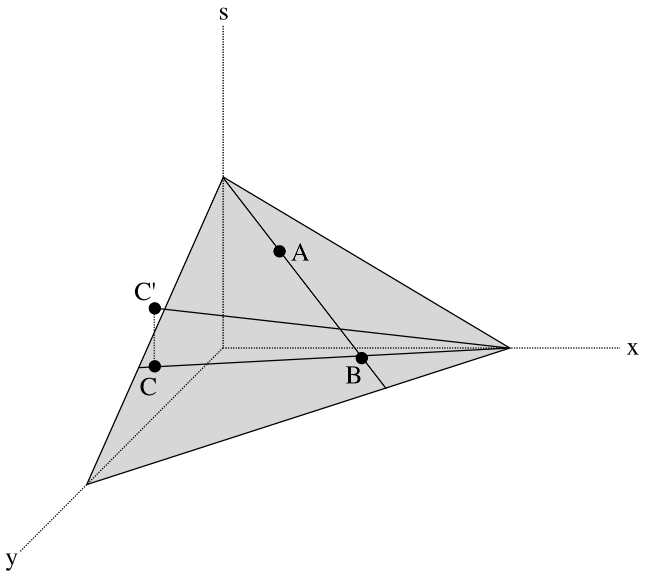


FIG. 4.—Example of numerical error in selecting point *C*. The *x*- and *y*-axes are the two main variables, and *s* is the slack variable introduced by the simplex linear programming algorithm. The gray region is the plane on which solutions lie. Point *C* lies far enough from point *B* that numerical error is introduced, leading to the selection of C', which lies outside the gray solution space. Subsequent similar sampling leads to exponentially fast growing error. The error in the diagram is exaggerated for clarity.

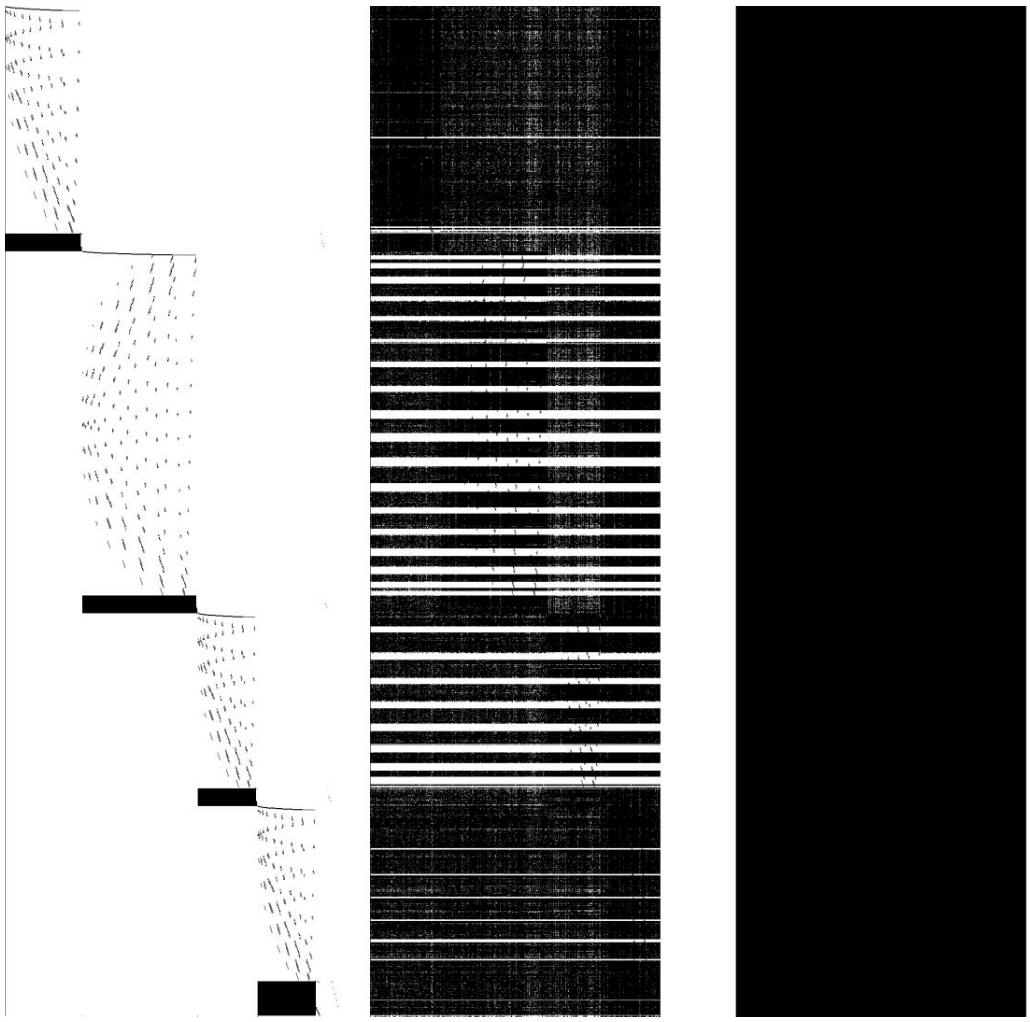


FIG. 5.— Three simplex tableaus displayed graphically. The left image is the tableau with the original constraints in place. Lens asymmetry can be seen in the block that is twice as tall as the other three. The middle image is after a feasible solution is found. The third image is after 200 models. Black represents nonzero values.

The source of the error can be seen in Figure 4. The figure has been exaggerated for clarity. Sample points are constrained to lie on the shaded surface. After sampling points A and B, point C is the next intended point, but because of the limits of machine accuracy, C' is taken instead. If the problem only occurred once, the error would be below the noise in the system, but each sample introduces more error because the next sample depends on the position of the previous sample.

Using the notation of Algorithm S, the further γ_{i+1} is chosen from γ_i, the larger the error. This is a simple lever; the error is proportional to (a/b), where $a = \gamma_{i+1} - \alpha_i$ and $b = \gamma_i - \alpha_i$. Successive errors are compounded over N iterations,

$$\epsilon = \prod_{i}^{N} (a_i/b_i). \tag{18}$$

Sampling uniformly along the line suppress the error, because points are chosen close to γ_i as often as far away. If $a_i \geq b_i$, ϵ grows without bound. If $\ln \epsilon > 0$, then $\langle a/b \rangle \geq N$, in which case the error is reported and the program halts.

A number of technical improvements were also made to the implementation of the simplex algorithm. As mentioned above, the simplex algorithm is used to find a new vertex in Algorithm S by maximizing an *objective function* subject to the linear constraints that form the simplex. Each iteration moves to a new vertex that increases the objective function until no further vertex can be found. The linear constraints are stored in a matrix called a *tableau*. The algorithm moves to the next vertex by rewriting the tableau—an operation known as a pivot. For very large problems, the pivot is the bottleneck. This work improves the performance by parallelizing the pivot on a shared-memory machine. For even larger problems than are faced here, it may be necessary to extend this to a distributed-memory cluster of machines.

A further improvement was an optimization of the data structure used to store the tableau. While the tableau is initially sparse and previous versions of PixeLens stored it as such, the tableau quickly becomes dense after only a few pivots (Fig. 5). Storing the tableau as a dense matrix yields a significant performance boost.

3. TESTING HUBBLE TIME RECOVERY

How well does Algorithm S predict the Hubble time? Two tests were performed.

First, a blind test similar to that in Williams & Saha (2000) is performed. Four quad lenses were crafted assuming a particular Hubble time that was unknown to the author. These were, in fact, the same lenses as in the aforementioned paper, but rescaled to a Hubble time of 13.9 Gyr. The time delays were perturbed slightly to simulate errors. The Hubble time was recovered using

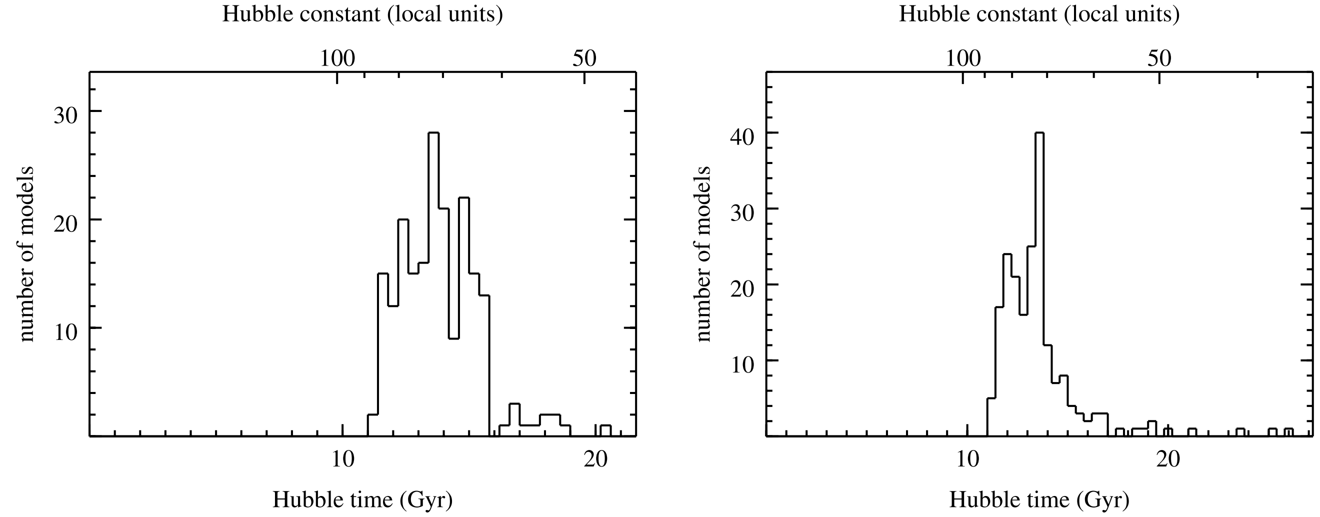


FIG. 6.—Two tests of the program. *Left*: Hubble time values recovered during a blind test. The lens was constructed by hand using an artificial value of the Hubble time (13.9 Gyr). *Right*: Time delays from a multiply lensed simulation galaxy with $H_0^{-1} = 14$ Gyr. There is a clear peak at 13.3 Gyr.

PixeLens and then the simulated Hubble time revealed. Figure 6 shows the histogram of Hubble times from 200 models. PixeLens predicts $H_0^{-1} = 13.7^{+1.5}_{-1.4}$ Gyr.

Second, five lenses, three doubles and two quads, were created by ray-tracing a galaxy from the *N*-body plus hydrodynamic simulation with $H_0^{-1} = 14$ Gyr described by Macciò et al. (2006). The galaxy is an E1 or E2 triaxial elliptical with about 80% dark matter. The histogram of Hubble times from 200 models is shown in the right panel in Figure 6. There is a clear peak with the predicted value at $H_0^{-1} = 13.3^{+1.4}_{-0.6}$ Gyr with 68% confidence. Within the errors, PixeLens successfully recovers the simulation Hubble time. Read et al. (2007) reconstruct the same lenses with a slightly different prior.

4. NEW 11 LENS RESULTS

With confidence founded in the results of § 3, an ensemble of lenses was modeled to find the true Hubble time. Saha et al. (2006) used 10 lenses[2] to constrain the Hubble time to $H_0^{-1} = 13.5^{+2.5}_{-1.3}$ Gyr. Subsequently, Vuissoz et al. (2007) have reported on a new time delay measurements for J1650+4251. Combining this new lens measurement with the 10 lenses used previously, all 11 lens were simultaneously modeled to predict tighter bounds on the Hubble time. The distribution of Hubble times is shown in Figure 7.

At 68% confidence, the new predicted value is

$$H_0^{-1} = 13.7^{+1.8}_{-1.0}\ \text{Gyr}\ \left(H_0 = 71^{+6}_{-8}\ \text{km s}^{-1}\ \text{Mpc}^{-1}\right).$$

Figure 8 shows the ensemble average of the mass and arrival time surface for J1650+4251 as recovered by PixeLens. Average mass maps for the other lenses are similar to those in Saha et al. (2006, Fig. 2).

To put this into context, the results of other techniques are listed below. The units are in H_0, which is found more often in the literature than H_0^{-1}. The latter appears more naturally in lensing, though, hence the presentation of the above estimates. The first set of errors are statistical and the second set (when applicable) are systematic. This list is summarized by the plot in Figure 9.

1. $H_0 = 73 \pm 3$ km s^{-1} Mpc^{-1} from the cosmic microwave background fluctuation spectrum (Spergel et al. 2007). The Hubble constant is just one value in a multiparameter fit.
2. $H_0 = 68 \pm 6 \pm 8$ km s^{-1} Mpc^{-1} using a different Monte Carlo method to combine lenses (Oguri 2007).
3. $H_0 = 62.3 \pm 1.3 \pm 5.0$ (Sandage et al. 2006) and $H_0 = 73 \pm 4 \pm 5$ (Riess et al. 2005) from Cepheid-calibrated luminosity of Type Ia supernovae. This is independent of the global geometry.
4. $H_0 = 66^{+11+9}_{-10-8}$ km s^{-1} Mpc^{-1} from the Sunyaev-Zel'dovich effect (Jones et al. 2005). As with lensing, a global geometry is assumed and the Hubble time is measured.
5. $H_0 = 72 \pm 8$ (Freedman et al. 2001) using a variety of Cepheid-calibrated indicators. This is again, independent of the global geometry.

In the future, better predictions may be obtained with improved priors and tighter constraints on galaxy structure. Simply

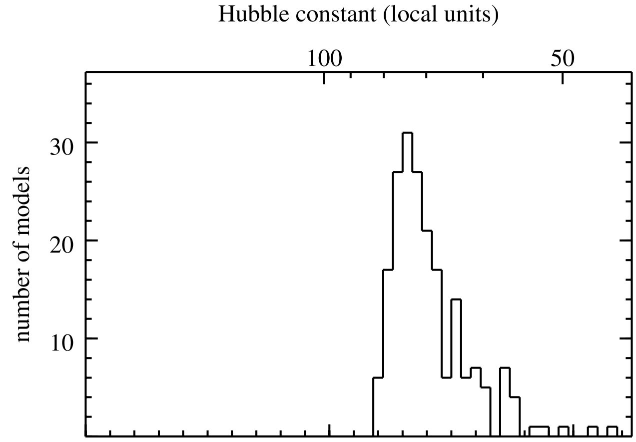


FIG. 7.—Hubble time values found from simultaneously modeling 11 lenses. The peak occurs at 13.7 Gyr.

---

[2] The 10 lenses are J0911+055, B1608+656, B1115+080, B0957+561, B1104−181, B1520+530, B2149−274, B1600+434, J0951+263, and B0218+357.

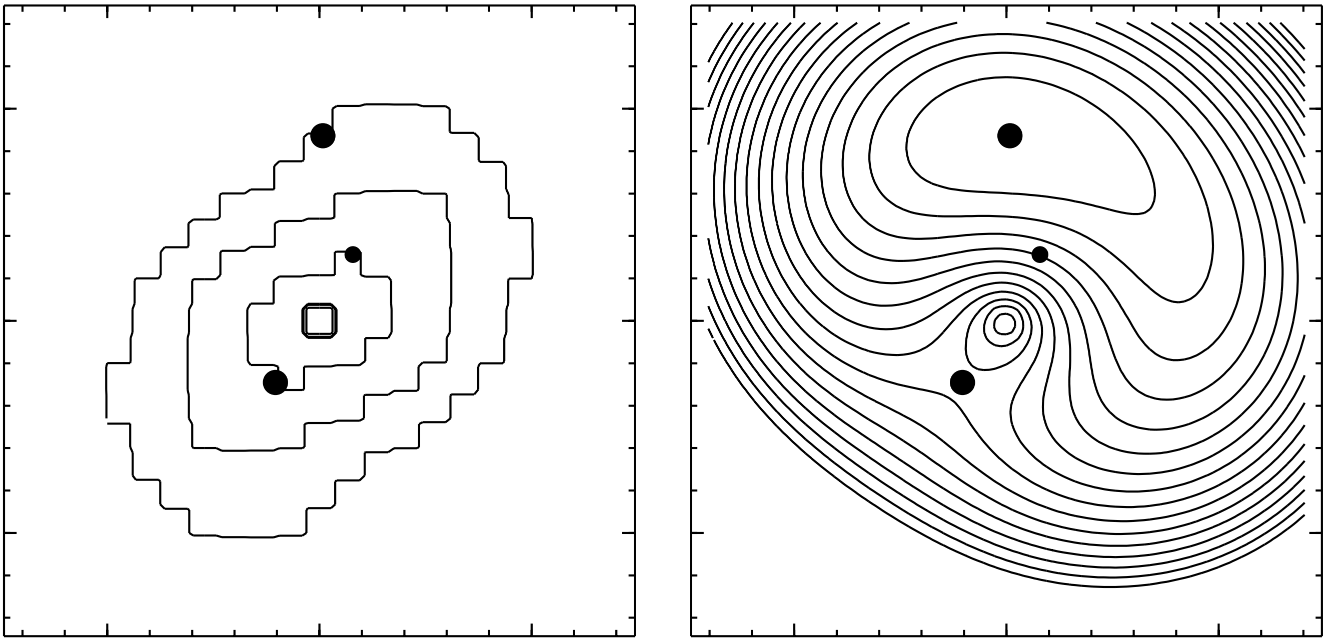


FIG. 8.—Example of the output from PixeLens: the ensemble average of the mass (*left*) and arrival time surface (*right*) of J1650+4251.

adding more lenses will also improve the predictive power of PixeLens, but may not help in understanding the different sources of degeneracies and developing better priors.

5. SUMMARY

Pixelated lens reconstruction is an example of free-form modeling. Such modeling has the advantage that one does not have to presuppose what the important parameters might be and can let the generated models be a guide to finding those parameters. Free-form modeling has many applications and was used early by Schwarzschild to show the existence of triaxial stellar systems in equilibrium.

Applied to gravitational lensing, the free-form models are implemented as pixelated models whereby the mass sheet of the lens is discretized into many small square pixels. The mass in each pixel is recovered using an MCMC technique using linear programming to probe the solutions which reconstruct the observed data. The software PixeLens produces an ensemble of hundreds of such models. The ensemble provides Bayesian statistics about the variety of possible lens reconstructions.

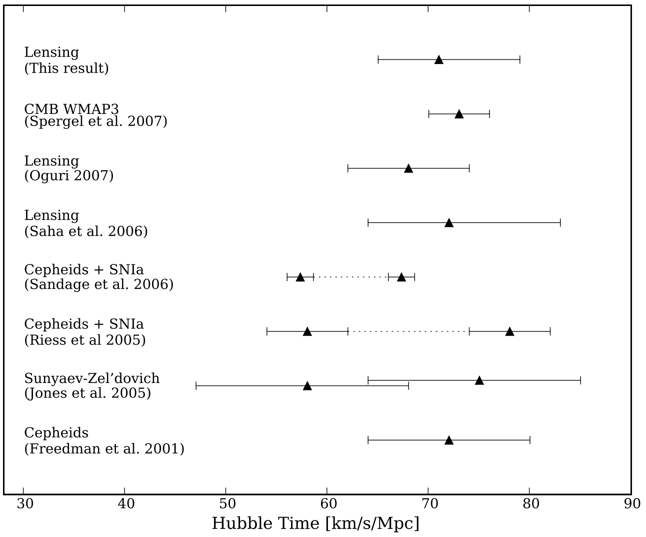


FIG. 9.—Recent Hubble time measurements from a variety of methods. Multiple error bars for a single reference are present when there are systematic errors in addition to statistical ones.

The constraints that define the mass models are explained. The linear constraints form a hyperdimensional solution space from which the models are drawn. The sampling algorithm has been improved over previous software versions, and although it was shown that the algorithm does not uniformly sample the solution space, it is argued that this is undesirable for this problem. The implementation was parallelized for multiprocessor, shared-memory machines. Future work will include controlling numerical round-off errors that will become significant with larger problems.

The new version of PixeLens was applied to an ensemble of 11 lenses to determine a new value for the Hubble time: $H_0^{-1} = 13.7^{+1.8}_{-1.0}$ Gyr within 68% confidence.

Further research into galaxy and cluster structure is needed to improve the priors. The estimates of galaxy morphology have been conservative, but tighter constraints will lead to better results. Furthermore, model ensemble building can be applied to other areas, even to the original problems of Schwarzschild.

Pixelated lens modeling is on the cutting edge of gravitational lens research, promising to provide great insight into the structure of galaxies, the distribution of dark matter, and the fundamental nature of the universe. But there are still many challenges both scientifically and computationally.

I would like to extend my sincere appreciation for the help I received with many aspects of this paper, in particular, Joachim Stadel for a critical insight concerning the material in § 2.2; Peter Englmaier, Tristen Hayfield, and Justin Read for the hours spent considering different sampling techniques; and Prasenjit Saha for patiently answering my many lensing questions and ever so subtly nudging me to finish. I would also like to thank the anonymous referee for useful comments and suggestions on making the paper clearer and more concise.

APPENDIX

PixeLens GRAVITATIONAL LENS MODELING SOFTWARE

PixeLens is freely available under the GNU General Purpose License. Source code is naturally included. The program is cross platform and an I/O limited version even runs in a web browser. The version used in this paper is v1.88.[3]

---

[3] For more information, visit http://www.qgd.uzh.ch.

REFERENCES

AbdelSalam, H. M., Saha, P., & Williams, L. L. R. 1998, AJ, 116, 1541

Bernstein, G., & Fischer, P. 1999, AJ, 118, 14

Binney, J. J., Dowrick, N. J., Fisher, A. J., & Newman, M. E. J. 1992, The Theory of Critical Phenomena (Oxford: Oxford Univ. Press)

Binney, J., Gerhard, O. E., Stark, A. A., Bally, J., & Uchida, K. I. 1991, MNRAS, 252, 210

Blandford, R. D., & Kundic, T. 1997, in The Extragalactic Distance Scale, ed. M. Livio, M. Donahue, & N. Panagia (Cambridge: Cambridge Univ. Press), 60

Blandford, R., & Narayan, R. 1986, ApJ, 310, 568

Bradač, M., Schneider, P., Lombardi, M., & Erben, T. 2005, A&A, 437, 39

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. 2001, Introduction to Algorithms (Cambridge: MIT Press)

Dantzig, G. B. 1963, Linear Programming and Extensions (Princeton: Princeton Univ. Press)

Diego, J. M., Protopapas, P., Sandvik, H. B., & Tegmark, M. 2005, MNRAS, 360, 477

Ferreras, I., Sakellariadou, M., & Furqaan, M. Y. 2008, Phys. Rev. Lett., 100, 031302

Freedman, W. L., et al. 2001, ApJ, 553, 47

Jones, M. E., et al. 2005, MNRAS, 357, 518

Keeton, C. R., & Winn, J. N. 2003, ApJ, 590, 39

Kochanek, C. S. 2004, in Gravitational Lensing: Strong, Weak, & Micro, ed. G. Meylan, P. Jetzer, & P. North (Berlin: Springer), 91

Macciò, A. V., Moore, B., Stadel, J., & Diemand, J. 2006, MNRAS, 366, 1529

Oguri, M. 2007, ApJ, 660, 1

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. 1986, Numerical Recipes: The Art of Scientific Computing (1st ed.; Cambridge: Cambridge Univ. Press)

Read, J. I., Saha, P., & Maccio, A. V. 2007, ApJ, 667, 645

Riess, A. G., et al. 2005, ApJ, 627, 579

Saha, P. 2003, Principles of Data Analysis (Great Malvern: Cappella Archive)

Saha, P., Coles, J., Macciò, A. V., & Williams, L. L. R. 2006, ApJ, 650, L17

Saha, P., & Williams, L. L. R. 1997, MNRAS, 292, 148

———. 2004, AJ, 127, 2604

———. 2006, ApJ, 653, 936

Sandage, A., Tammann, G. A., Saha, A., Reindl, B., Macchetto, F. D., & Panagia, N. 2006, ApJ, 653, 843

Schwarzschild, M. 1979, ApJ, 232, 236

Spergel, D. N., et al. 2007, ApJS, 170, 377

Trotter, C. S., Winn, J. N., & Hewitt, J. N. 2000, ApJ, 535, 671

Vuissoz, C., et al. 2007, A&A, 464, 845

Wambsganss, J., & Paczynski, B. 1994, AJ, 108, 1156

Williams, L. L. R., & Saha, P. 2000, AJ, 119, 439

Young, P., Gunn, J. E., Oke, J. B., Westphal, J. A., & Kristian, J. 1981, ApJ, 244, 736

4 Paper III

Weak Microlensing

...whereby the Universe learns Tai χ^2

Mon. Not. R. Astron. Soc. **000**, 000–000 (0000) Printed 10 November 2009 (MN LaTeX style file v2.2)

Weak Microlensing

Jonathan Coles[1]*, Prasenjit Saha[1], and Hans Martin Schmid[2]
[1] *Institut für Theoretische Physik, Universität Zürich, Winterthurerstr. 190, 8006 Zürich, Switzerland*
[2] *Institut für Astronomie, ETH Zürich, 8093 Zürich, Switzerland*

10 November 2009

ABSTRACT

A nearby star having a near-transit of a galaxy will cause a time-dependent weak lensing of the galaxy. Because the effect is small, we refer to this as weak microlensing. This could provide a useful method to weigh low-mass stars and brown dwarfs. We examine the feasibility of measuring masses in this way and we find that a star causes measurable weak microlensing in a galaxy even at 10 Einstein radii away. Of order one magnitude $I \leqslant 25$ galaxy comes close enough to one or other of the ~ 100 nearest stars per year.

Key words: gravitational lensing; stars: low-mass, brown dwarfs; methods: statistical

1 INTRODUCTION

Gravitational lensing by Galactic stars has come a long way since the low-probability assessment of Einstein (1936). The first detections (Alcock et al. 1993; Aubourg et al. 1993; Udalski et al. 1993) have been followed by $\sim$ 4000 more, including a few detections of planets around the main lensing star (e.g., Gould et al. 2006).

Yet despite the now abundant examples of microlensing, accurate measurements of the lensing mass are still rare. The reason is that a microlensing light-curve on its own tells us precisely how long proper motion takes to traverse an Einstein radius, but the physical length r_E of the Einstein radius and its angular size θ_E remain unknown. Hence, only a broad statistical statement about the stellar mass can be made in most cases.

Refsdal (1966) anticipated the problem and suggested a way to overcome it: If both stars are directly observable, their relative proper motion would (together with the light curve) supply θ_E, while a second observatory in the solar system would provide r_E. Several related ideas have been forwarded, notably by An et al. (2002), who exploited a combination of caustic-crossing times and finite source size effects to obtain the first microlensing mass measurement. A similar strategy has recently been used by Gould et al. (2009) to measure the mass of a brown dwarf to 10%.

Another possibility is to measure the lensing effect of a star on a *galaxy.* Paczyński (1996) pointed out that as a Galactic star makes a near-transit of a galaxy, the latter will undergo a small shift in its apparent centroid. For this variety of microlensing to be observable, θ_E needs to be larger than the precision of image centroiding–although it can be below resolution–and proper motions need to be large. Nearby stars are the only realistic prospect, since

$$\theta_E \approx 0\overset{\prime\prime}{.}09 \left(\frac{M/M_\odot}{D_\text{lens}/\text{pc}}\right)^{1/2} \tag{1}$$

for sources at infinity, and proper motions are $\sim 1''$ per year. Microlensing by nearby stars would have none of the degeneracy problems mentioned above; the lens being at known distance, the mass is the only unknown parameter. Moreover, such events are predictable long in advance. The transit may be on the order of a few months allowing for observations of the galaxy both before and during the lensing event. Ideally, the star would not transit directly across the galaxy since the star must be masked out for proper observations of the galaxy.

Paczyński (1996) suggested that microlensing centroid shifts could be used to measure down to masses of nearby brown dwarfs. In this paper, we suggest a refinement of Paczyński's idea which could make it much more effective. Rather than just the galaxy centroid shift, the whole weakly-lensed image of the galaxy could be exploited to infer the lensing mass. We develop a technique to extract the weak-lensing effects and estimate the mass of the star.

2 A FITTING METHOD

Consider a star with Einstein radius θ_E whose sky position at epoch t is z_t. Lensing by this star maps a point θ in the image plane to a point ϕ in the source plane such that

$$\phi = \theta - \theta_E^2 \frac{(\theta - z_t)}{|\theta - z_t|^2} \tag{2}$$

where z_t, θ, ϕ are two-dimensional vectors.

Next, we consider a galaxy whose unlensed brightness distribution S is expanded in terms of basis functions $B^n(\theta)$

---

* E-mail: jonathan@physik.uzh.ch

as

$$S(\theta) = \sum_n a^n B^n(\theta) \qquad (3)$$

where a^n are the expansion coefficients. In the presence of lensing, the brightness distribution of the galaxy will be

$$S_t(\theta) = \sum_n a^n B^n(\phi(\theta, z_t, \theta_E)) \quad , \qquad (4)$$

where S now has a time dependence due to the position of the star z_t. Since lensing conserves surface brightness, the lensed surface brightness at θ equals the unlensed surface brightness at ϕ.

Pixelating the image plane, we write the pixel-wise brightness distribution as

$$D_{t,ij} \equiv \sum_n a^n L^n_{t,ij} \qquad (5)$$

where $L^n_{t,ij}$ represents a lensed and pixelated basis function

$$L^n_{t,ij} \equiv B^n(\phi(\theta_{ij}, z_t, \theta_E)) \quad . \qquad (6)$$

The expression in Equation 5 is our model for the lensed brightness distribution. If the observed distribution is $d_{t,ij}$ then the likelihood is

$$\mathcal{L}(a^n, \theta_E) \propto \prod_{t,ij} \exp\left[-\tfrac{1}{2}\,\sigma^{-2}_{t,ij}\,(d_{t,ij} - D_{t,ij})^2\right] \qquad (7)$$

where $\sigma_{t,ij}$ is the pixel-wise noise, which we assume is Gaussian. The model has a complicated dependence on θ_E, but only a linear dependence on the expansion coefficients a^n. Since we are not especially interested in the a^n, we marginalise them out by standard methods (for example, chapter 5 of Saha 2003). The marginalised likelihood is given by

$$2\ln\mathcal{L}(\theta_E) = \ln|\det C| + \sum_{mn} P_m P_n C_{mn} - \sum_{t,ij} \sigma^{-2}_{t,ij}\, d^2_{t,ij} \quad . \qquad (8)$$

Here

$$P_n \equiv \sum_{t,ij} \sigma^{-2}_{t,ij}\, d_{t,ij}\, L^n_{t,ij} \qquad (9)$$

represents a kind of projection of the data on the model, while

$$C^{-1}_{mn} \equiv \sum_{t,ij} \sigma^{-2}_{t,ij}\, L^m_{t,ij}\, L^n_{t,ij} \qquad (10)$$

is the inverse covariance matrix. We can relate $\mathcal{L}(\theta_E)$ to an effective χ^2 as just

$$\mathcal{L}(\theta_E) = \exp(-\chi^2/2) \quad . \qquad (11)$$

We are now prepared to estimate the mass of a star by its lensing effect on a background galaxy. Given the pixel-wise brightness $d_{t,ij}$ and noise level $\sigma_{t,ij}$, we simply need to calculate the effective χ^2 and minimize with respect to θ_E.

As an aside, $|\det C|$ in Equation 8 will typically overflow floating-point arithmetic, while $\ln|\det C|$ will fit quite nicely. Hence we compute an alternative form, namely, $\ln|\det C| = \sum \ln\lambda_n$, where λ_n is the nth eigenvalue of C.

$\gamma_{\rm tot}$	$\frac{1}{2}$,1,3,7,12	$\times 10^6$
$N_{\rm obs}$	$2, 3, 4, 5$	
$\theta_{E,{\rm true}}$	$10, 20, 30, 40, 50$	[milliarcsec]
p	0.08, 0.126, 0.232, 0.454, 0.903	[arcsec]

Table 1. Parameters used for simulated observations: $\gamma_{\rm tot}$ is the total number of photons over all observation epochs (including one unlensed observation), $N_{\rm obs}$ is the number of epochs, $\theta_{E,{\rm true}}$ is the Einstein radius we wish to recover, and p is the closest approach of the star on the sky plane to the background galaxy.

3 SIMULATED OBSERVATIONS

We tested the above scheme using simulated data, focusing specifically on the dependence on four quantities: the actual Einstein radius of the star $\theta_{E,{\rm true}}$, the closest approach or impact parameter p, the number of observed epochs (or t values) $N_{\rm obs}$, and the total number of photons collected $\gamma_{\rm tot}$. The values are summarized in Table 1. The full matrix of these parameters was tested, giving a total of 500 simulated observational programs.

For the exact form of the unlensed surface brightness in Equation 3 we chose

$$S(\theta) = \exp\left(-7.67\left[\sqrt{(\theta/R_e)^2 + R_c^2}\right]^{1/4}\right) \qquad (12)$$

with $R_e = 1$. This is simply a de Vaucouleurs profile modified by a core radius R_c of 2 pixels to mimic the effect of the telescope PSF on a singular cusp.

We then considered 71×71 pixels imaging a patch of sky $2''$ on a side and centred on the galaxy. With this resolution each pixel is about $l = 0.''028$ across, which, for example, would be equivalent to about one pixel per resolution element of the Nasmyth Adaptive Optics System (NAOS) Near-Infrared Imager and Spectrograph (CONICA) camera (Lenzen et al. 2003; Rousset et al. 2003). At a redshift of $z \sim 0.5$, where we expect to find most of our background galaxies, $1''$ is about 6 kpc. The corresponding lensed pixelated brightness is

$$S_{t,ij} = S(\phi(\theta_{ij}, z_t, \theta_{E,{\rm true}})) \quad . \qquad (13)$$

We normalized this brightness to the total number of photons

$$\sum_{t,ij} S_{t,ij} = \gamma_{\rm tot}, \qquad (14)$$

and thus $\gamma_{\rm tot}/N_{\rm obs}$ is the number of photons per image. Considering the normalized $S_{t,ij}$ as the expectation value of the pixel-wise photon count, we then drew the simulated data $d_{t,ij}$ from a Poisson distribution. The noise level $\sigma_{t,ij}$ was taken as $\sqrt{d_{t,ij}}$, which greatly simplifies Equations 8 and 9.

In practice, the star will be masked out during observations. To simulate this, we discarded the pixels within one pixel-width of the star.

As basis functions we chose two-dimensional Hermite functions or shapelets (e.g., Refregier 2003). The scale parameter for the shapelets was set to $\beta = 0.''2$ and we used $N_{\rm basis} = 20 \times 20$ basis functions in all. These settings allow structures as large as $\theta_{\rm max} = \beta(2\sqrt{N_{\rm basis}} + 1)^{1/2} = 1.''28$ and as small as $\theta_{\rm min} = \beta(2\sqrt{N_{\rm basis}} + 1)^{-1/2} = 0.''03$ to be resolved (Melchior et al. 2007).

For each of the 500 test cases, we considered one epoch

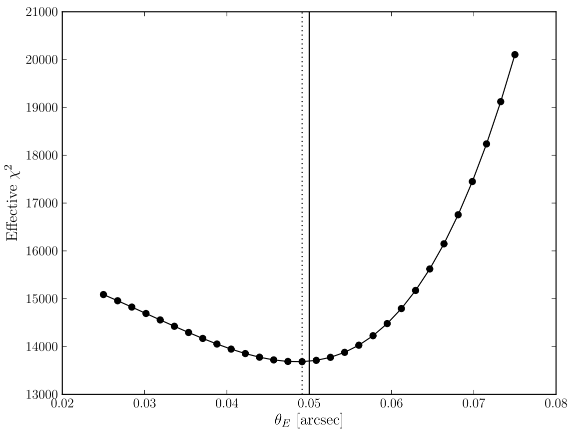


Figure 1. Plot of the effective χ^2 as a function of θ_E, for a simulated observation program with $N_{\rm obs} = 2$ epochs, total photons $\gamma_{\rm tot} = 3 \times 10^6$ collected, $\theta_{E,\rm true} = 0\farcs05$ and impact parameter $p = 0\farcs126$. The solid vertical line marks $\theta_{E,\rm true}$ and the dashed line marks the best fit. The formal count of degrees of freedom is 2 epochs $\times\, 71^2$ pixels $-\, 20^2$ basis functions $= 9682$, so the effective reduced $\chi^2 \approx 1.4$. The dependence of the fit on θ_E is non-linear, hence the asymmetric shape of the curve.

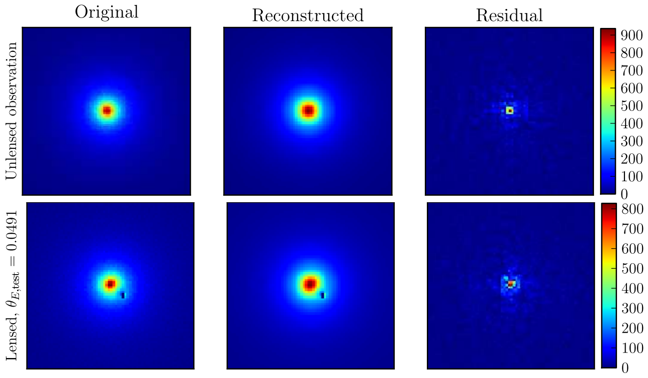


Figure 2. Details of the best-fit θ_E for the simulated observation program of Figure 1. The upper row refers to an unlensed epoch $t = 0$ and the lower row to the lensed epoch $t = 1$. The left column shows the simulated data $d_{t,ij}$, the middle column shows the best reconstruction $D_{t,ij}$, while the right column is $|d_{t,ij} - D_{t,ij}|$. The scale is in units of photons. The star has been masked out as would be done during an observation.

$t = 0$ with the galaxy unlensed and additional epochs $t > 0$ with the star at $z_t = (5lt, -l2^i \mid i \in [1, \ldots, 5])$. The selection of i corresponds to the selection of p: A large p implies a large i. This choice of position puts the star at the centre of a pixel. Extensive testing has shown a sensitivity to placing stars near pixel edges, whereby recovery of the data is degraded if the star is too close to pixel boundaries.

In Figure 1 we show the effective χ^2 as a function of θ_E for one of the simulated data sets. In Figure 2 we show the simulated images, along with the reconstructed and residual images for the best-fit θ_E. Examining such plots is a good indicator of potential problems. For example, if too few basis functions are used, a grid-like pattern shows up in the reconstructions and the residual, and recovery of θ_E is very poor.

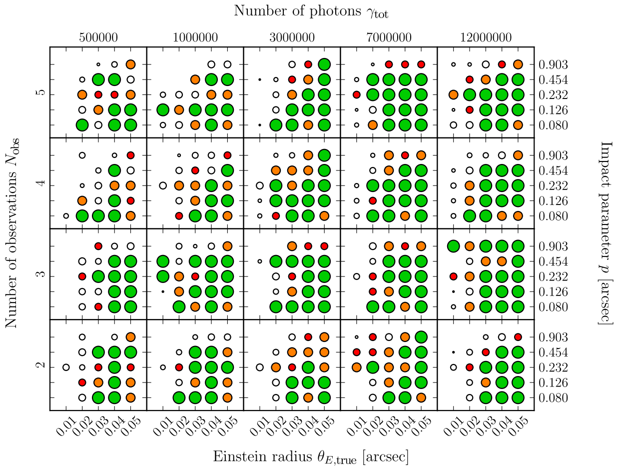


Figure 3. Mass uncertainty as a function of *four* quantities: θ_E, impact parameter p, total galaxy photons collected $\gamma_{\rm tot}$, and the number of epochs $N_{\rm obs}$. Within each box, θ_E and p are varied at fixed $\gamma_{\rm tot}, N_{\rm obs}$. The latter two quantities are varied between boxes, as labelled. Circles indicate the error in mass (i.e., in θ_E^2): Large, green circles denote $< 5\%$ error, orange circles 5–11% error, red circles 11-20% error, and small, open circles are used if the mass error was $> 20\%$. Missing circles mean that no mass estimate could be made. Of all the tests, 26% have errors less than 5% and 39% have errors less than 11%. Of those tests with filled circles, 52% have errors less than 5% and 79% have errors less than 11%.

Figure 3 summarizes the complete suite of 500 simulated observation programs, showing the mass-recovery errors as a function of $\theta_{E,\rm true}$, p, $\gamma_{\rm tot}$ and $N_{\rm obs}$. The following conclusions can be easily read off:

- The mass range of nearby brown dwarfs is accessible, since θ_E down to $0\farcs02$ can be measured with the resolution considered.
- Impact parameters of $p \sim 10\,\theta_E$ are small enough, but if p is too small the galaxy can be obscured by the star mask, leading to poor results.
- Of order a million photons from the galaxy are needed, and a few times this are desirable, but it does not matter much whether these are concentrated in two epochs or distributed among several epochs.

4 EVENT RATES

We now consider how likely is it to find a star crossing near a background galaxy. For this analysis we used the Research Consortium on Nearby Stars (RECONS) list of the 100 nearest stellar systems (Henry 2009). The proper motions and estimated masses of the stars in these systems are plotted in Figure 4. In Figure 5 we have plotted the area of sky swept out by Einstein radii per year, or $2\,\mu\theta_E$ where μ is the proper motion.

If we restrict ourselves to masses $< 0.5 M_\odot$ and proper motions $> 0\farcs5\,/{\rm yr}$, we are left with 85 stars. Assuming, as seen in our tests, that a galaxy within $10\,\theta_E$ is a candidate, we sum $20\,\mu\theta_E$ over these stars. The total available area is $\sim 70\,{\rm arcsec}^2$ per year.

The GalaxyCount program (Ellis & Bland-Hawthorn 2007) estimates $\simeq 1$ galaxy with magnitude $I \leqslant 25$ within

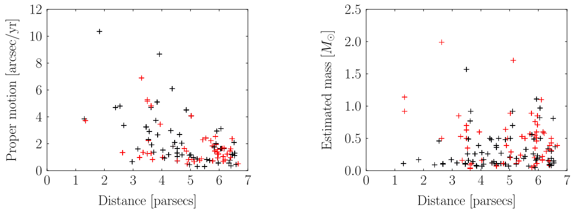


Figure 4. The nearest 100 stellar systems (139 stars) from the RECONS catalogue. Shown are the proper motions (left) and estimated masses (right) for the sample. Systems in red have been excluded from the discussion in Section 4. The five highest mass stars are, α Centauri A+B, Sirius, Procyon, and Altair.

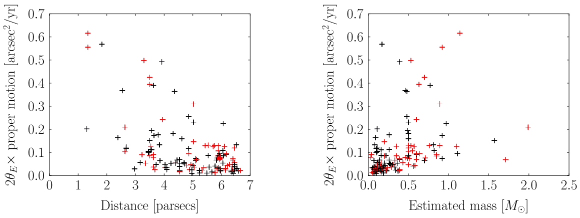


Figure 5. Sky area coming within one Einstein radius of each nearby star per year, plotted against distance (left), and against estimated stellar mass (right). Systems in red have been excluded from the discussion in Section 4.

a sky area of 70 arcsec^2. This provides a rough estimate of the rate of observable weak microlensing events.

5 OBSERVATIONAL PROSPECTS

Observing the weak lensing of a faint galaxy by a nearby star would require a high resolution ($\approx 0.''05$) imager with high contrast capabilities. A $0.5M_\odot$ star at 5 pc has a brightness of $I \approx 6.5$mag while $I \approx 12$mag for a $0.1M_\odot$ star at the same distance. Thus a contrast in the range $\Delta I = 12-18$mag must be achieved at a separation of about $0.''05$ for typical events. This is quite a challenge for existing instruments. Fortunately, rapid progress can be expected in this field by instruments currently built for the imaging of planetary systems with 8m-10m telescopes and further significant progress will be possible with extremely large telescopes and high contrast imagers in space. They will provide very high contrast observations $\Delta m > 20$mag and allow mass determinations of many nearby stars using weak microlensing of faint background galaxies as advocated in this Letter. Any background light that is not from the galaxy can still be considered part of the source as it will either be lensed or remain relatively constant throughout the duration of the complete observation program. An 8m class telescope with 30-50% efficiency collects about 50,000 photons/hr. Thus, a typical program might need between 20 and 100 hours to expect reasonable results.

With existing instruments it should already be possible to observe weak microlensing in favourable cases where the impact parameter is small and the optical resolution is higher than that considered here. Nearby ($d \approx 5$ pc) brown dwarfs with a mass of $\approx 0.05M_\odot$ ($\theta_E \approx 0.''01$) such as SCR 1845-6357 B (at 3.9 pc), DENIS 0255-4700 (5.0 pc), 2MASS 0415-0935 (5.7 pc), or GJ 570 D (5.9 pc), have $I \approx 17-20$ mag and $V \approx 22-25$ mag and they are not or not much brighter than the abundant backgound galaxies. Low mass stars and substellar objects are red or extremely red and imaging observations of blue star-forming galaxy at short wavelengths is favoured because the image contamination of the lensed galaxy by the PSF of the lensing object is strongly reduced. It seems that HST or an adaptive optics systems (e.g. with laser guide star) at a large telescope working at short wavelengths $< 1\mu$m should be capable of achieving successful observations for certain weak microlensing events.

REFERENCES

Alcock C., Akerlof C. W., Allsman R. A., et al. 1993, Nature, 365, 621
An J. H., Albrow M. D., Beaulieu J.-P., et al. 2002, ApJ, 572, 521
Aubourg E., Bareyre P., Bréhin S., et al. 1993, Nature, 365, 623
Einstein A., 1936, Science, 84, 506
Ellis S. C., Bland-Hawthorn J., 2007, MNRAS, 377, 815
Gould A., Udalski A., An D., et al. 2006, ApJ, 644, L37
Gould A., Udalski A., Monard B., et al 2009, ApJ, 698, L147
Henry T. J., , 2009, RECONS, `http://joy.chara.gsu.edu/RECONS/`
Lenzen R., Hartung M., Brandner W., et al. 2003, in M. Iye & A. F. M. Moorwood ed., Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4841 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, NAOS-CONICA first on sky results in a variety of observing modes. pp 944–952
Melchior P., Meneghetti M., Bartelmann M., 2007, A&A, 463, 1215
Paczyński B., 1996, Acta Astronomica, 46, 291
Refregier A., 2003, MNRAS, 338, 35
Refsdal S., 1966, MNRAS, 134, 315
Rousset G., Lacombe F., Puget P., et al. 2003, in P. L. Wizinowich & D. Bonaccini ed., Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4839 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, NAOS, the first AO system of the VLT: on-sky performance. pp 140–149
Saha P., 2003, Principles of Data Analysis. Great Malvern, UK, Cappella Archive, 2003
Udalski A., Szymanski M., Kaluzny J., Kubiak M., Krzeminski W., Mateo M., Preston G. W., Paczynski B., 1993, Acta Astronomica, 43, 289

5 Paper IV

COSMOGRAIL: The COSmological MOnitoring of GRAvItational Lenses VII: Time delays and the Hubble constant from WFI J2003-4723

> ...whereby the Universe learns to tell time but is not so confident about its age.

A&A 488, 481–490 (2008)
DOI: 10.1051/0004-6361:200809866

Astronomy
&
Astrophysics

COSMOGRAIL: the COSmological MOnitoring of GRAvItational Lenses[⋆,⋆⋆]

VII. Time delays and the Hubble constant from WFI J2033–4723

C. Vuissoz[1], F. Courbin[1], D. Sluse[1], G. Meylan[1], V. Chantry[2,⋆⋆⋆], E. Eulaers[2], C. Morgan[3,4], M. E. Eyler[4],
C. S. Kochanek[3], J. Coles[5], P. Saha[5], P. Magain[2], and E. E. Falco[6]

[1] Laboratoire d'Astrophysique, École Polytechnique Fédérale de Lausanne (EPFL), Observatoire de Sauverny, 1290 Versoix, Switzerland
e-mail: `christel.vuissoz@epfl.ch`
[2] Institut d'Astrophysique et de Géophysique, Université de Liège, Allée du 6 août 17, Sart-Tilman, Bât. B5C, 4000 Liège, Belgium
[3] Department of Astronomy and the Center for Cosmology and Astroparticle Physics, The Ohio State University, Columbus, OH 43210, USA
[4] Department of Physics, United States Naval Academy, 572C Holloway Road, Annapolis MD 21402, USA
[5] Institute of Theoretical Physics, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
[6] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge MA 02138, USA

Received 28 March 2008 / Accepted 2 July 2008

ABSTRACT

Gravitationally lensed quasars can be used to map the mass distribution in lensing galaxies and to estimate the Hubble constant H_0 by measuring the time delays between the quasar images. Here we report the measurement of two independent time delays in the quadruply imaged quasar WFI J2033−4723 ($z = 1.66$). Our data consist of R-band images obtained with the Swiss 1.2 m EULER telescope located at La Silla and with the 1.3 m SMARTS telescope located at Cerro Tololo. The light curves have 218 independent epochs spanning 3 full years of monitoring between March 2004 and May 2007, with a mean temporal sampling of one observation every 4th day. We measure the time delays using three different techniques, and we obtain $\Delta t_{B-A} = 35.5 \pm 1.4$ days (3.8%) and $\Delta t_{B-C} = 62.6^{+4.1}_{-2.3}$ days $(^{+6.5\%}_{-3.7\%})$, where A is a composite of the close, merging image pair. After correcting for the time delays, we find R-band flux ratios of $F_A/F_B = 2.88 \pm 0.04$, $F_A/F_C = 3.38 \pm 0.06$, and $F_{A1}/F_{A2} = 1.37 \pm 0.05$ with no evidence for microlensing variability over a time scale of three years. However, these flux ratios do not agree with those measured in the quasar emission lines, suggesting that longer term microlensing is present. Our estimate of H_0 agrees with the concordance value: non-parametric modeling of the lensing galaxy predicts $H_0 = 67^{+13}_{-10}$ km s^{-1} Mpc^{-1}, while the Single Isothermal Sphere model yields $H_0 = 63^{+7}_{-3}$ km s^{-1} Mpc^{-1} (68% confidence level). More complex lens models using a composite de Vaucouleurs plus NFW galaxy mass profile show twisting of the mass isocontours in the lensing galaxy, as do the non-parametric models. As all models also require a significant external shear, this suggests that the lens is a member of the group of galaxies seen in field of view of WFI J2033−4723.

Key words. gravitational lensing – cosmology: cosmological parameters – quasars: individual: WFI J2033−4723

1. Introduction

When a quasar is gravitationally lensed and we observe multiple images of the source there are light travel time differences between the images. Any intrinsic variation of the quasar is observed at different times in each image with a measurable "time delay" between them. Refsdal (1964) first showed that time delays provide a means of determining the Hubble constant H_0 independent of any local distance calibrator, provided a mass model can be inferred for the lensing galaxy. Conversely, one can also assume H_0 in order to study the distribution of the total mass in the lensing galaxy.

During the past 25 years, time delays have been measured in only 17 systems, at various accuracy levels (see Oguri 2007, for a review). As the error in the time delay propagates directly into H_0, it is important to make it as small as possible. Unfortunately, most existing time delays have uncertainties of the order of 10% that are comparable to the current uncertainties in H_0. This uncertainty can be reduced by increasing the sample of lenses with known time delays, and by simultaneously fitting all lenses in the sample with a common value for H_0 (Saha et al. 2006b,a; Coles 2008).

COSMOGRAIL is an optical monitoring campaign that aims at measuring time delays for a large number of gravitationally lensed quasars to accuracies of a few percent using a network of 1- and 2-m class telescopes. The first result of this campaign was the measurement of the time delay in the doubly imaged quasar SDSS J1650+4251 to an accuracy of 3.8% based on two

---

[⋆] Based on observations obtained with the 1.2 m EULER Swiss Telescope, the 1.3 m Small and Moderate Aperture Research Telescope System (SMARTS) which is operated by the SMARTS Consortium, and the NASA/ESA *Hubble Space Telescope* as part of program HST-GO-9744 of the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555.

[⋆⋆] Table 4 is only available in electronic form at the CDS via anonymous ftp to `cdsarc.u-strasbg.fr (130.79.128.5)` or via `http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/A+A/488/481`

[⋆⋆⋆] Research Fellow, Belgian National Fund for Scientific Research (FNRS).

Article published by EDP Sciences

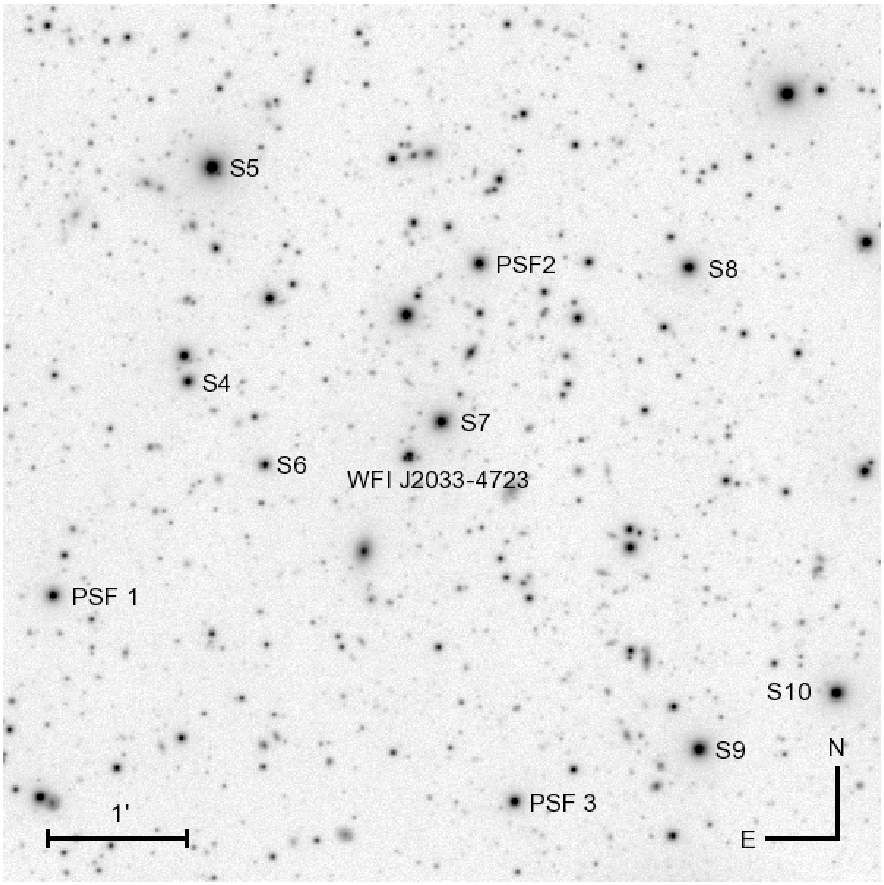


Fig. 1. The 6.3′ × 6.3′ field of view around WFI J2033−4723. This image is the central part of a combination of 418 *R*-band frames obtained with the 1.2 m EULER Telescope with a total exposure time of 48 h and a mean seeing of 1.3″. The three stars PSF1−3 used to model the Point Spread Function (PSF) and the 7 reference stars S4−10 used for frame alignment and flux calibration are indicated.

observing seasons of data (Vuissoz et al. 2007). COSMOGRAIL complements a second monitoring group whose most recent results are a delay for HE 1104-1805 (Poindexter et al. 2007). In this paper we present time-delay measurements for the quadruply imaged quasar WFI J2033−4723 based on merging 3 years of optical monitoring data from the two groups. In a companion effort, Morgan et al. (2008) analyzed the merged data for the two-image lens QJ0158−4325, succeeding in measuring the size of the source accretion disk but failing to measure a time delay due to the high amplitude of the microlensing variability in this system.

WFI J2033−4723 ($20^{\mathrm{h}}33^{\mathrm{m}}42\overset{\mathrm{s}}{.}08$, $-47°23'43\overset{''}{.}0$; J2000.0) was discovered by Morgan et al. (2004) and consists of 4 images of a $z = 1.66$ quasar with a maximum separation of 2.5″. The lens galaxy was identified by Morgan et al. (2004) and has a spectroscopic redshift of $z_{\mathrm{lens}} = 0.661 \pm 0.001$ (Eigenbrod et al. 2006). The lens appears to be the member of a group, with at least 6 galaxies within 20″ of the lens (Morgan et al. 2004), and we will have to account for this environment in any lens model.

We describe the monitoring data and its analysis in Sect. 2. In Sect. 3 we present the near-IR *Hubble Space Telescope* (HST) observations that we used to obtain accurate differential astrometry of the lens components and surface photometry of the lens galaxy. We estimate the time delays in Sect. 4 and model them using parametric (Sect. 5) and non-parametric (Sect. 6) models for the mass distribution of the lens galaxy. We summarize our results in Sect. 7.

2. Photometric monitoring

Our 3-year photometric monitoring of WFI J2033−4723 was carried out from March 2004 to May 2007 with the 1.2 m EULER telescope and the 1.3 m SMARTS telescope located in Chile at La Silla and the Cerro Tololo Interamerican Observatory (CTIO), respectively. WFI J2033−4723 was monitored from both sites for three years during its visibility window from early March to mid-December.

The 1.2 m EULER telescope is equipped with a 2048 × 2048 CCD camera which has 0.344″ pixels and produces an image with an 11′ field of view. The mean sampling of the EULER data is one epoch every five days, where each epoch consists of five dithered 360 s images taken with an *R*-band filter. The worst gaps due to weather or technical problems are 2−3 weeks. The EULER data set consists of 141 epochs of data obtained between May 2004 and May 2007. The image quality varies between 0.9″ and 2.0″ *FWHM* over 3 years, with a median of 1.4″.

The 1.3 m SMARTS telescope is equipped with the dual-beam ANDICAM (DePoy et al. 2003) camera. Here we use only the optical channel which has 0.369″ pixels and a 6.5′ × 6.3′ field of view. The mean sampling of the SMARTS data is one epoch every eight days, with three 300 s exposures at each epoch. The SMARTS data set consists of 77 epochs of data obtained between March 2004 and December 2006. The seeing on the images varies between 0.5″ and 2.0″, with a median of 1.4″.

The combined data set consists of 218 observing epochs comprising 956 images covering the common field of view shown in Fig. 1. The average temporal sampling when

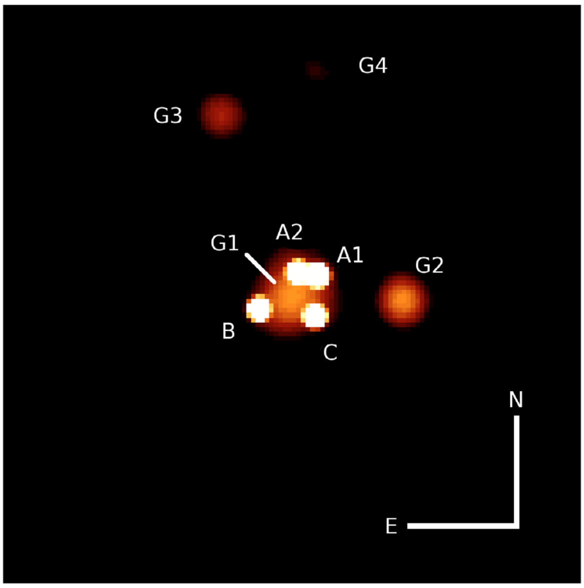


Fig. 2. Result of the simultaneous deconvolution of the 956 R-band images (EULER+SMARTS) of WFI J2033−4723. The pixel size of this image is half the pixel size of the EULER detector, i.e., 0.172″, and the resolution is 2 pixels Full-Width-Half-Maximum, i.e., 0.344″. The field of view is 22″ on a side. Two galaxies G2 and G3 are seen to the West and North of the main lensing galaxy G1. G3 is part of a group included in the lens modeling, while G1 and G2 are modeled individually (see Sect. 5).

WFI J2033−4723 was visible is one observation every 4 days over a period of three years, one of the best sets of monitoring data available for a lensed quasar.

The EULER data are reduced using the automated pipeline described in Vuissoz et al. (2007) and the SMARTS data with the SMARTS pipeline, using standard methods. The reduced frames are then aligned and interpolated to a common reference frame, one of the best-seeing (1″) EULER images, taken on the night of 5 April 2006. The 10 stars (PSF1−3 and S4−10) shown in Fig. 1 are used to determine the geometric transformation needed for each EULER and SMARTS image to match the reference frame. The transformation includes image parity, rotation, shifting and rescaling. These 10 stars are also used to determine the photometric zero point of each image relative to the reference image. After interpolation, cosmic rays are removed using the L.A.Cosmic algorithm (van Dokkum 2001). We check that no data pixels are mistakenly removed by this process.

The light curves of the quasars are measured by simultaneously deconvolving all the images using the MCS algorithm (Magain et al. 1998). This method has already been successfully applied to the monitoring data of several lensed quasars (e.g. Vuissoz et al. 2007; Hjorth et al. 2002; Burud et al. 2002a,b). The deconvolved images have a pixel scale of 0.172″ (one-half the pixel scale of the EULER data) and are constructed to have a Gaussian PSF with a 2 pixel (0.344″) *FWHM*. The Point Spread Function (PSF) for each of the 956 images is constructed from the three stars PSF1−3 (see Fig. 1). During the deconvolution process, the relative positions of the quasar images are held fixed to the values derived from fitting the HST images in Sect. 3, while their fluxes are allowed to vary from frame to frame. The flux, position and shape of the lensing galaxy are the same for all frames, but the values vary freely as part of the fit.

Figure 2 shows an example of a deconvolved image. It is clear that we will have no difficulty estimating the fluxes of components B and C separately. Components A_1 and A_2, however, are separated by only 0.724″, which is only twice the resolution of our deconvolved images, and remain partially blended after deconvolution. Since the delay between these images should be very small, we will sum the fluxes of the two images and consider only the light curve of the total flux $A = A_1 + A_2$. The resulting R-band light curves are displayed in Fig. 3.

We also display in Fig. 3 the deconvolved light curve of the isolated star $S6$, which has roughly the same color as WFI J2033−4723. Each point is the mean of the images taken at a given epoch and the error bar is the 1σ standard error of the mean. The light curve is flat, with a standard deviation over the 3 years of monitoring of $\sigma_{\rm tot} = 0.010$ mag about its average, which is roughly consistent with the mean error bar of $\sigma_{\rm mean} = 0.006$ mag of the individual epochs.

The dispersion of the points in the star's light curve reflects both statistical errors and systematic errors from the photometric calibrations and the construction of the PSF. To the extent that all the light curves suffer from the same systematic errors, we can correct the quasar's light curves by subtracting the residuals of the star's light curve from each of them. We then define the uncertainty in a quasar's light curve as the quadrature sum of the uncertainties in the two light curves. This procedure will increase the photon noise but should minimize the systematic errors.

3. HST near-IR imaging

We determine the relative positions of the lens components and the light profile for the main lens galaxy G1 and its closest neighbor G2 (see Fig. 2) by analyzing our HST images of the system. These data were obtained in the framework of the CASTLES program (Cfa-Arizona Space Telescope LEns Survey), which provides HST images for all known gravitationally lensed quasars. We deconvolve these images using a modified version of the MCS deconvolution algorithm for images with poor knowledge of the PSF (Magain et al. 2007). We previously used this approach to unveil the faint Einstein ring and the lensing galaxy hidden in the glare of the quasar images of the so-called "cloverleaf" HE1413+117 (Chantry & Magain 2007). We analyze the Near Infrared Camera and Multi-Object Spectrometer (NICMOS) $F160W$ (H-band) images obtained with the NIC2 camera. The data consist of four dithered MULTIACCUM images with a total exposure time of 2752 s and a mean pixel scale of 0.07568″ (Krist & Hook 2004). We calibrate the images using CALNICA, the HST image reduction pipeline, subtract constants for each quadrant of NIC2 to give each image a mean background of zero, and create a noise map for each frame.

The images are simultaneously deconvolved in a manner similar to that used for the EULER and SMARTS data. The NICMOS frames lack any PSF stars, so we construct the PSF using the quasar images themselves in the iterative method of Chantry & Magain (2007). We first estimate the PSF of each frame using Tiny Tim (Krist & Hook 2004) and then we deconvolve them to have the final Gaussian PSF. During the deconvolution, each image is decomposed into a set of point sources and any extended flux. The latter is then reconvolved to the resolution of the original data and subtracted from the four initial frames, leading to images with far less contamination by extended structures. Four new PSFs are estimated from these new images, and we carry out a new deconvolution. The process is repeated until the residual maps are close to being consistent with the estimated noise (e.g. Courbin et al. 1998). In this case,

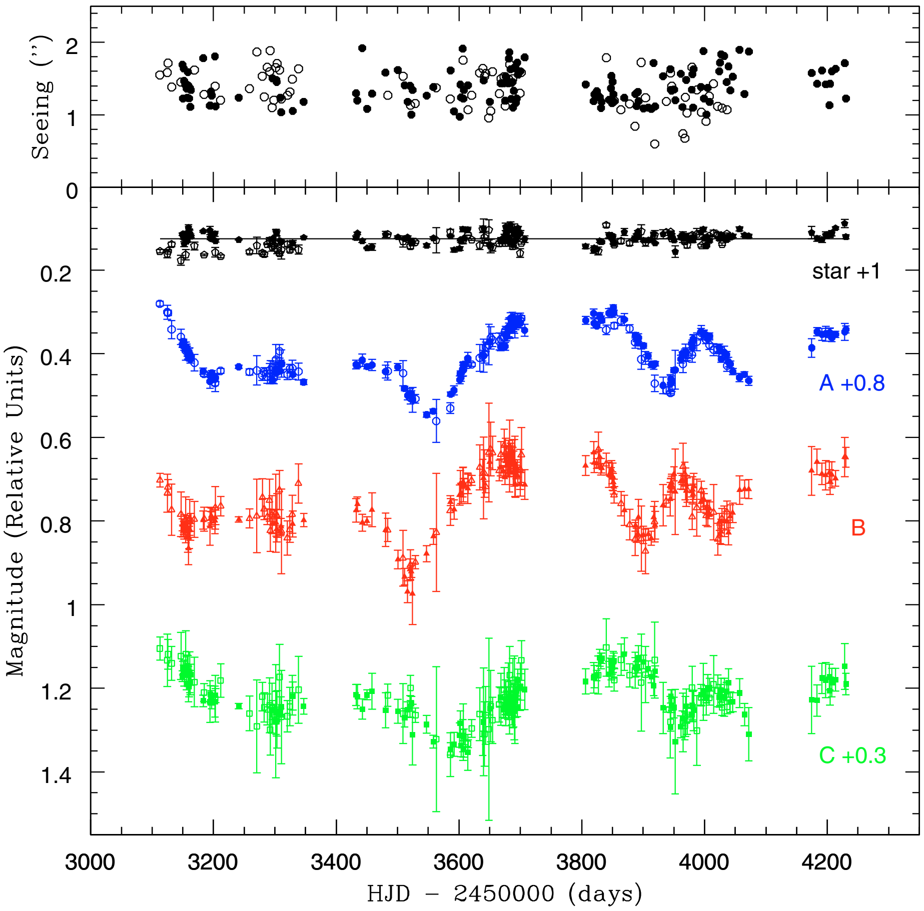


Fig. 3. Our *R*-band light curves obtained for WFI J2033−4723 as well as for the reference star S6 (see Fig. 1). The magnitudes are given in relative units. The filled symbols correspond to the EULER observations while the SMARTS data points are marked by open symbols. Components A_1 and A_2 were summed into one single component *A*. The curves have been shifted in magnitude for visibility, as indicated in the figure.

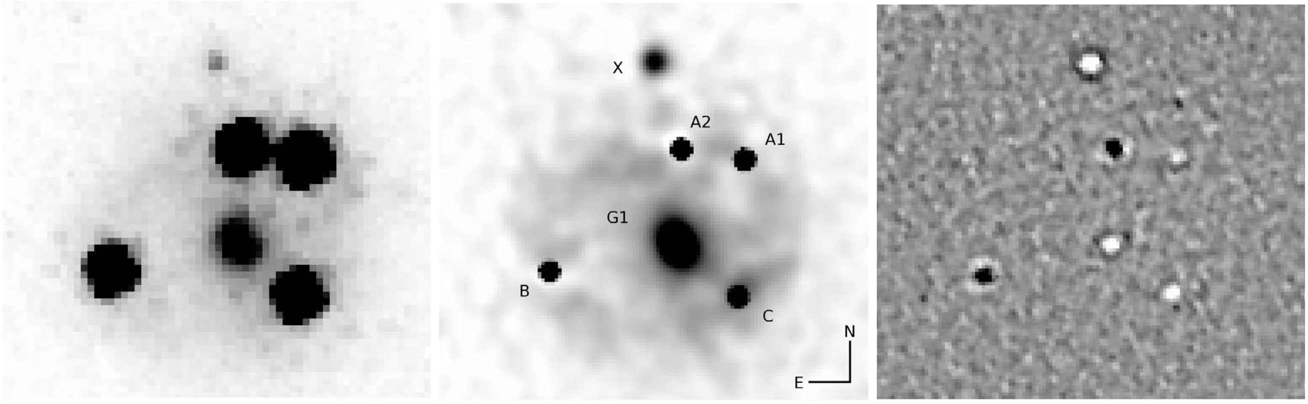


Fig. 4. *Left*: deep NICMOS2 image, taken in the *F*160*W*-band. This image is a combination of 4 frames, for a total exposure time of 46 min. North is up and East to the left. The field of view is 4″ on a side. *Middle*: simultaneous deconvolution of the individual NICMOS images (see text), using the MCS deconvolution algorithm. The PSF in this image is an analytical Gaussian with 2 pixels Full-Width-at-Half-Maximum (*FWHM*), i.e., the resolution is 0.075″. The pixel size is 0.035″, i.e., oversampled by a factor of two compared to the original pixel size. *Right*: residual map of the deconvolution, with the cut levels set to ±5σ. Only minor structures are seen in the center of the sharpest objects, which is acceptable given the quality of the NICMOS PSF.

convergence is reached after three iterations and the final reduced χ^2 is 3.59. The final deconvolved image shown in Fig. 4 has half the pixel scale of the initial images and a Gaussian PSF with a *FWHM* of 0.075″.

As part of the MCS deconvolution we also fit analytical models to the main lens galaxy (G1) and its nearby companion G2. The main lens is an early-type galaxy (Eigenbrod et al. 2006), as its companion is likely to be, so we use elliptical de Vaucouleurs profiles for both. The uncertainties are estimated by the scatter of the measurements from a separate set of fits to each independent image. We also estimate that there are systematic errors in the astrometry from the NICMOS pixel scale and focal plane distortions of order 2 milli-arcsec based on our earlier fits to the NICMOS data of H1413+117 (Chantry & Magain 2007). These

Table 1. Relative astrometry and photometry for the four components of WFI J2033−4723 and for the lensing galaxies G1 and G2.

ID	$\Delta\alpha$ ($''$)	$\Delta\delta$ ($''$)	Magnitude
B	0.	0.	17.77 ± 0.02
$A1$	−2.1946 ± 0.0004	1.2601 ± 0.0003	17.16 ± 0.02
$A2$	−1.4809 ± 0.0004	1.3756 ± 0.0005	17.52 ± 0.02
C	−2.1128 ± 0.0003	−0.2778 ± 0.0003	17.88 ± 0.02
G1	−1.4388 ± 0.0019	0.3113 ± 0.0008	18.59 ± 0.03
G2	−5.4100 ± 0.0006	0.2850 ± 0.0003	18.14 ± 0.02

Table 2. Shape parameters for the main and secondary lensing galaxies.

Obj.	PA (°)	Ellipticity	a_0 ($''$)	b_0 ($''$)
G1	27.8 ± 4.3	0.18 ± 0.03	0.665 ± 0.036	0.556 ± 0.025
G2	6.4 ± 3.1	0.15 ± 0.02	0.389 ± 0.004	0.334 ± 0.005

systematic errors are compatible with the Lehár et al. (2000) comparison of NICMOS and VLBI astrometry for radio lenses.

The relative astrometry and photometry of the lens components and of the lensing galaxies G1 and G2 are given in Table 1. Coordinates are relative to image B, in the same orientation as Fig. 4. The photometry is in the Vega system. For each measurement, we give the 1σ internal error bars, to which a systematic error of 2 milli-arcsec should be added. The models for G1 and G2 are presented in Table 2, with the effective semi-major and semi-minor axes of the light distribution a_0 and b_0. Each measurement is accompanied by its 1σ error bar.

4. Time delay measurement

We measure the time delays between the blended light curve of A_1/A_2 and images B and C using three different techniques: *i* the minimum dispersion method of Pelt et al. (1996); *ii* the polynomial method of Kochanek et al. (2006); and *iii* the method of Burud et al. (2001). Since WFI J2033−4723 shows well-defined variations (see Fig. 3), it is already clear by visual inspection that $\Delta t_{B-A} \sim 35$ days and $\Delta t_{B-C} \sim 65$ days.

4.1. Minimum dispersion method

In the minimum dispersion method, time delays are computed for pairs of light curves using a cross-correlation technique that takes into account irregular sampling. The two light curves are first normalized to have zero mean. Then, one of the light curves is used as a reference and the second curve is shifted relative to it by a range of time delays. For each delay, we calculate the mean dispersion between the shifted points and their nearest temporal neighbors in the reference light curve. The best time-delay estimate is the one that minimizes this dispersion function. Since the mean sampling of our curves is one epoch every four days and since there is a limit to the number of time delays that can be tested independently, we test time delays in steps of 2 days. Figure 5 shows an example of a dispersion curve where we have then fit a parabola and set the best time delay to be the one corresponding to the minimum of the parabola.

There is, however, a complication in the step of normalizing the light curves, arising from sampling the light curve of each lensed image over a different time period of the intrinsic source variability (Vuissoz et al. 2007). We solve this problem by computing the dispersions as a function of a small magnitude shift Δm in the normalization, measuring both the minimum dispersion $D_{\min}(\Delta m)$ and the best fitting time delay $\Delta t(\Delta m)$ as shown in Fig. 6. Our final value for the delay is the one corresponding to the shift Δm that minimizes the overall dispersion.

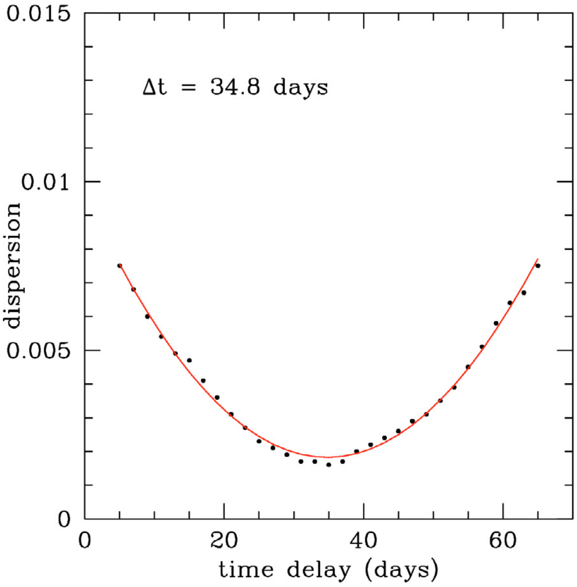


Fig. 5. Example of a dispersion curve as obtained from the minimum dispersion method, for components B and A. The position of the parabola minimum gives the time delay. Each point is separated by 2 days, i.e. about half the data mean sampling. The time delay indicated here is for only one realization of the boot-strap procedure (see text).

We then estimate the uncertainties by randomly perturbing the data points, based on a Gaussian distribution with the width of the measurement errors, and computing the dispersions and time delays again. We define the uncertainties by the 1σ dispersion in the results for 100 000 trials (Vuissoz et al. 2007). The resulting uncertainty estimates are symmetric about the mean, so our final estimates based on this method are

$$\Delta t_{B-A} = 35.6 \pm 1.3 \text{ days } (3.6\%)$$
$$\Delta t_{B-C} = 64.6 \pm 3.4 \text{ days } (5.3\%). \quad (1)$$

While we have not taken microlensing effects into account for this analysis, it should matter little, as the method is not very sensitive (Eigenbrod et al. 2005) to the very low amplitude microlensing variability observed for this system (see Sect. 4.2).

4.2. Polynomial fit of the light curves

In the polynomial method (Kochanek et al. 2006), the intrinsic light curve of the source is modeled as a polynomial that is simultaneously fit to all three light curves. Each quasar image has an additional low order polynomial to model slow, uncorrelated variability due to microlensing. We increase the source polynomial order for each season until the improvement in the χ^2 statistics of the fits cease to be significant. This results in using polynomial orders of 11, 10, 17 and 4 for the four seasons of data. The low amplitude microlensing variations are modeled with a simple linear function for the four seasons. Figure 7 shows the best fits to the data and the residuals from the model. The effects of microlensing in this system are very small, with variations of only ~0.01 mag over three years. As with the minimum dispersion method, we estimate the uncertainties by randomly

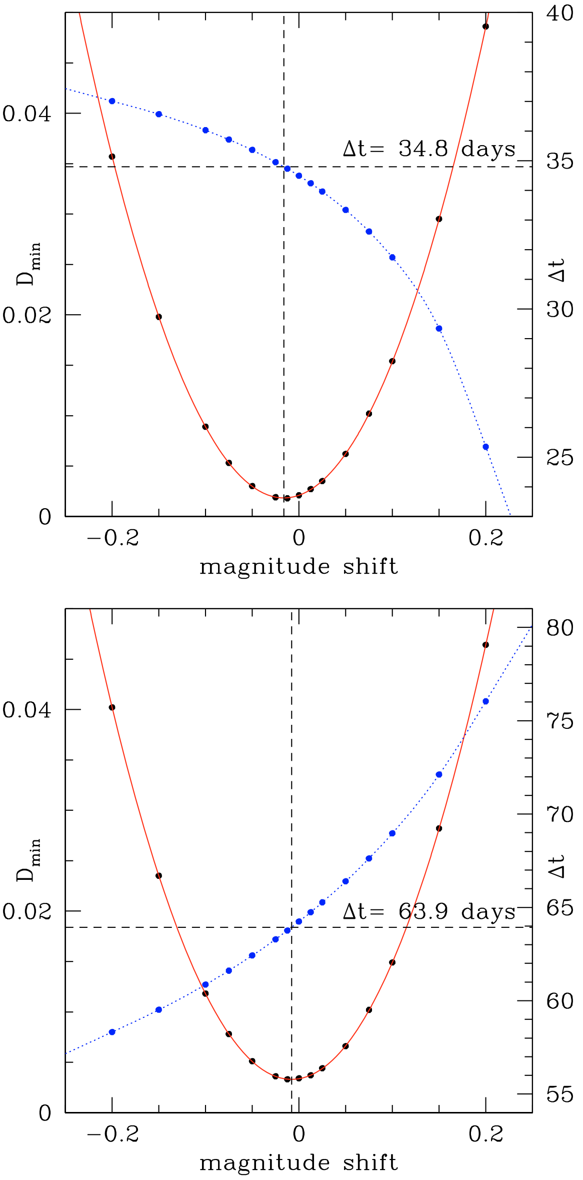


Fig. 6. Variation of the dispersion function minimum $D_{\rm min}$ (red solid curves), as a function of the magnitude shift used for the normalization (see Sect. 4.1). Each $D_{\rm min}$ corresponds to a different estimate of the time delay, indicated on the blue dotted curves. The final time delay is the one with the lowest $D_{\rm min}$. *The top panel* is for the *B-A* time delay, *the bottom* one for *B-C*. Time delays indicated here are for only one realization of the boot-strap procedure (see text).

perturbing the light curves 100 000 times and using the standard deviation of the trials as the error estimates to find that

$$\begin{aligned} \Delta t_{B-A} &= 35.0 \pm 1.1 \;\; \text{days} \;\; (3.0\%) \\ \Delta t_{B-C} &= 61.2 \pm 1.5 \;\; \text{days} \;\; (2.4\%). \end{aligned} \quad (2)$$

4.3. Numerical modeling of the light curves

Our last approach is based on the method described in Burud et al. (2001), which determines the time delay between a pair of light curves using a gridded numerical model for the source light curve. For a series of time delays, we fit the data with a flux ratio between the two curves, and a linear trend for microlensing on each full light curve. The numerical source model is smoothed on the scale of the temporal sampling, based on a smoothing function weighted by a Lagrange multiplier. The best time delay is the one that minimizes the χ^2 between the model and data points.

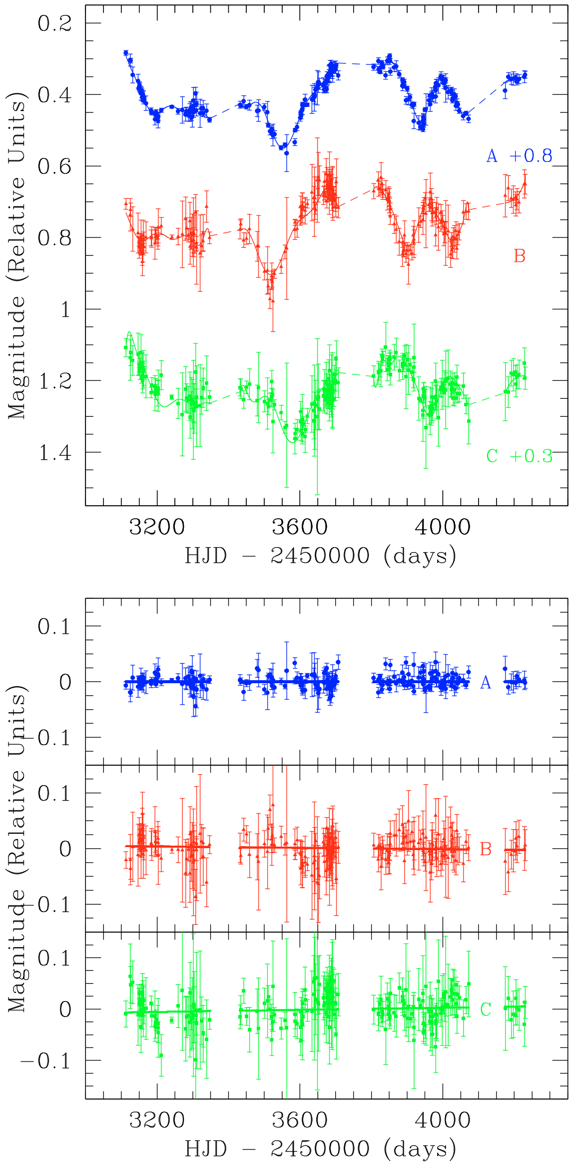
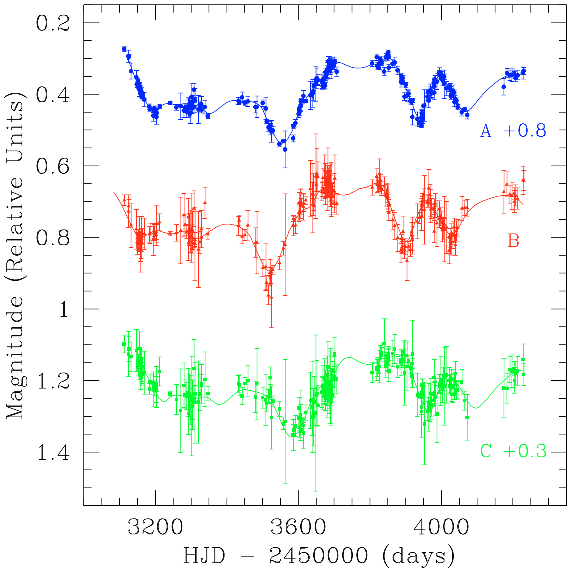


Fig. 7. *Top*: best polynomial fit to the light curves, which are shifted vertically for display purpose. *Bottom*: the residuals of the fit.

This method has several advantages: first, none of the data light curves is taken as a reference: they are all treated on an equal basis. Furthermore, as the model is purely numerical, no assumption is made on the shape of the quasar's intrinsic light curve (except that it is sufficiently smooth). Finally, a model light curve is obtained for the intrinsic variations of the quasar, as for the polynomial fit method (see Sect. 4.2).

We have applied this method to the two pairs of light curves of WFI J2033−4723. The resulting fits to the light curves and their residuals are shown in Fig. 8. Using a Monte Carlo method to estimate the uncertainties, we find from 7000 trials (adding normally distributed random errors with the appropriate standard deviation on each data point):

$$\begin{aligned} \Delta t_{B-A} &= 36.0 \pm 1.5 \;\; \text{days} \;\; (4.2\%) \\ \Delta t_{B-C} &= 61.9\,^{+6.7}_{-0.5} \;\; \text{days} \;\; (^{+11\%}_{-1\%}). \end{aligned} \quad (3)$$



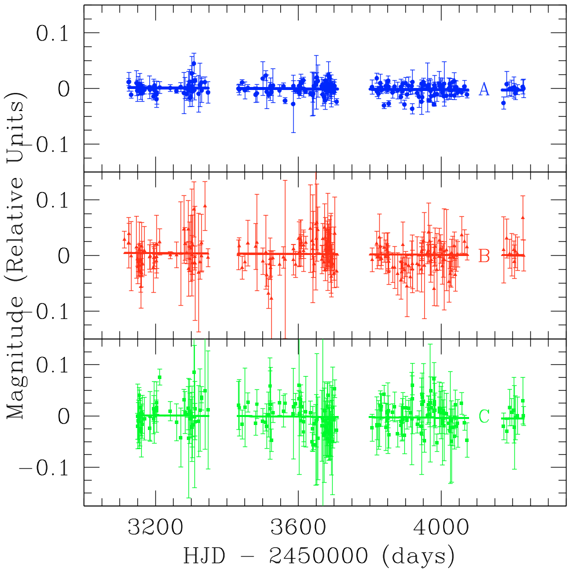


Fig. 8. *Top*: the light curves of the quasar images shown along with the best numerical model. *Bottom*: the residuals of the fit.

We note a secondary peak in the Δt_{B-C} Monte Carlo distribution, around 69 days, in addition to the main peak at 61.9 days. There is, however, no evidence of such a secondary peak in the results of the minimum dispersion method and the polynomial fitting technique. This translates into an asymmetrical error bar on the result obtained with the numerical fitting of the light curves, and is taken into account in our final estimate of the time delay between quasar images B and C.

4.4. Final time delays

For the final delay estimate we adopt the unweighted mean of the three methods, and we take as uncertainties the quadrature sum of the average statistical error and the dispersion of the results from the individual methods about their mean. Our final estimate of the time delays is

$$\Delta t_{B-A} = 35.5 \pm 1.4 \text{ days } (3.8\%)$$
$$\Delta t_{B-C} = 62.6\,^{+4.1}_{-2.3} \text{ days } (^{+6.5\%}_{-3.7\%}). \qquad (4)$$

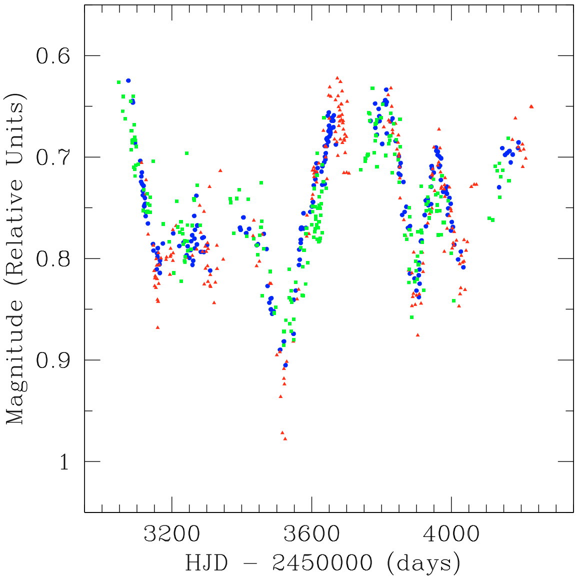


Fig. 9. Light curves of the three quasar images, shifted by their respective time delay and flux ratio. The blue circles correspond to image A, the red triangles to B and the green squares to C.

We cannot measure the time delay between the individual A_1 and A_2 light curves, but values larger than $\Delta t_{A1-A2} = 2$ days are incompatible with any of the models we consider in the following section. We can nevertheless estimate the flux ratio between A_1 and A_2. After correcting for the time delays, we find that the R-band flux ratios between the images are

$$\frac{F_A}{F_B} = 2.88 \pm 0.04, \quad \frac{F_A}{F_C} = 3.38 \pm 0.06, \quad \frac{F_{A1}}{F_{A2}} = 1.37 \pm 0.05. \quad (5)$$

Figure 9 displays the light curves obtained for the three quasar images, after shifting by the time delays and flux ratios. Note that these flux ratios differ from those measured by Morgan et al. (2004) from the MgII broad emission lines, probably due to long-term microlensing (on a longer scale than our monitoring 3-year baseline), as discussed in the next section.

5. Parametric modeling

5.1. Observational constraints

We constrain the mass models of WFI J2033−4723 using the positions of the four lensed images, the position of the lens galaxy G1 and the two delay measurements, for a total of 12 constraints. Except when indicated, we do not use the image flux ratios because they can be affected by extinction (Falco et al. 1999; Jean & Surdej 1998) and milli-lensing by substructures (Mao & Schneider 1998). We can also constrain the structure of G1 given its ellipticity e, position angle θ_e and effective radius R_e to the extent that these properties are correlated with those of its dark matter halo. Although a possible mismatch between the light and mass distributions is not impossible, we adopt the formal errors of 0.002″ on the position of the lens G1 (Table 2). This is motivated by the small offset between the centers of mass and light found by Yoo et al. (2006) in a sample of four lensing galaxies.

Finally, WFI J2033−4723 is located in a group of galaxies, labeled G2–G6 in Fig. 10. We include G2 as a singular isothermal sphere (SIS) in all our models since it is close (4″) and of similar luminosity to G1. As Morgan et al. (2004), we are unable to successfully model the system without including G2.

Table 3. Result of the parametric lens modeling.

Name	Comp.	Mass	e, θ_e	γ, θ_γ	#d.o.f.	χ^2	h	Comments
SIE+γ		$b = 0.96$	0.21, (20.8)	0.187, 7.4	1	15.3	–	Time delays not used
SIE+γ		$b = 0.94$	0.13, (30.5)	0.063, 24.6	2	3.6	$0.79^{+0.04}_{-0.02}$	$G_{\rm group}$ included
SIE+γ		$b = 0.97$	0.16, 84.4	0.059, 46.9	1	0.30	$0.63^{+0.07}_{-0.03}$	$G_{\rm group}$ included
dVC+γ		$b = 2.71$	0.20, (20.1)	0.305, 9.7	1	34.4	–	Time delays not used
dVC+γ		$b = 2.83$	0.18, 83.1	0.116, 64.5	1	0.01	$0.92^{+0.06}_{-0.03}$	$G_{\rm group}$ included
NFW+γ		$\kappa_{\rm s} = 0.20$	0.16, (27.4)	0.070, –3.6	1	0.38	–	$r_{\rm s} = 10''$ (fixed); time delays not used
NFW+γ		$\kappa_{\rm s} = 0.21$	0.15, 85.7	0.079, 9.5	1	0.06	$0.29^{+0.03}_{-0.03}$	$r_{\rm s} = 10''$ (fixed); $G_{\rm group}$ included
NFW+γ		$\kappa_{\rm s} = 0.09$	0.15, 85.4	0.076, 30.6	1	0.01	$0.63^{+0.10}_{-0.08}$	$r_{\rm s} = 1''$ (fixed); $G_{\rm group}$ included
dVC+NFW+γ	Light	$b = 1.56$	(0.17), (29.3)	0.057, 37.0	–	–	–	$R_{\rm e} = 0.608''$(fixed)
	Halo	$\kappa_{\rm s} = 0.082$	0.065, (29.3)		1	6.33	$0.78^{+0.12}_{-0.10}$	$r_{\rm s} = 10''$ (fixed); $G_{\rm group}$ included
dVC+NFW+γ	Light	$b = 1.53$	(0.16), (26.4)	0.075, 27.5	–	–	–	same model as above, with
	Halo	$\kappa_{\rm s} = 0.10$	0.43, 89.8		3	3.2	$0.69^{+0.20}_{-0.10}$	flux ratios included

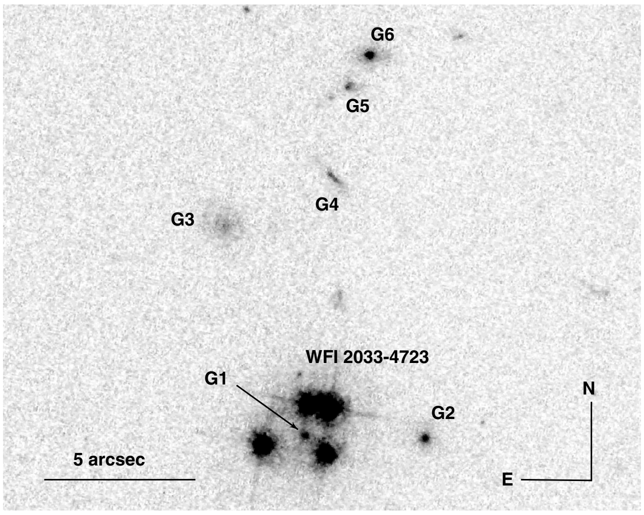


Fig. 10. Environment of WFI J2033−4723 as seen in an HST/ACS F814 (I-band) image. The main lens galaxy G1 and the close companion G2 were included in our analysis of the NICMOS image, and here we have labeled additional group members as G3−G6.

When enough observational constraints are available we further add galaxies G3−G6 as a SIS mass distribution located at the barycenter $G_{\rm group}$ of the group. In all models we include an external shear of amplitude γ and position angle θ_γ that represents the gravitational perturbations due to mass unaccounted for explicitly. We also experiment with including mass at the position of object X (Fig. 4, $2''$ North of G1) and find that doing so does not improve the models.

We consider a sequence of standard mass models for G1, including a singular isothermal ellipsoid (SIE), a de Vaucouleurs (dVC) model and a NFW model (Navarro et al. 1997), and we fit the data using `LENSMODEL` (v1.99g) (Keeton 2001). The results are summarized in Table 3, where Cols. #1 and #2 describe the model family, and #3 the mass parameter (either the Einstein radius b in arcseconds or the mean mass surface density $\kappa_{\rm s}$, as defined in Keeton 2001). Column #4 is for the ellipticity e and PA θ_e of the lens G1. Note that the measured PA of G1 is $\theta = 27.8°$ (Table 2). Column #5 gives the external shear amplitude γ and PA θ_γ, #6 the number of degrees of freedom for each model, and #7 the resulting reduced χ^2. Column #8 finally shows the best estimate for $h = H_0/100$. A minus sign in this column means that time delays are not used as constraints. All angles are given positive East of North, and values given in parentheses are fitted to the observations. All models assume $\Delta t_{A1-A2} = 2$ days and include the companion galaxy G2, with a resulting mass $0.1\, m_{\rm G1} < m_{\rm G2} < m_{\rm G1}$.

5.2. SIE models

Our first model consists of an SIE for G1, an SIS for G2 and an external shear. When we fit only the image positions but include a prior on the position angle θ_e (from Table 1) we do not find a good fit unless the constraint on the position of the lensing galaxy is relaxed. The prior on the position angle is justified by statistical studies finding correlations between the position angles but not the axis ratios of the visible and total mass distributions (Ferreras et al. 2008; Keeton et al. 1997). With the inclusion of the time delays we have enough constraints to add the group halo to the model. With the position angle of G1 constrained, we obtain poor fits to the data with reduced $\chi^2 = 3.6$ for $N_{\rm d.o.f.} = 2$. When we leave the position angle free, we find good fits but the model PA is 55° from the observed. These models have Hubble parameters of $h \simeq 0.63^{+0.07}_{-0.03}$ with the spread dominated by the degeneracies between the ellipticity and the shear.

5.3. De Vaucouleurs models

Next we consider a constant mass-to-light ratio model of the lens galaxy based on a De Vaucouleurs model. The position angle and the effective radius $R_{\rm e} = 0.608''$ (the geometric mean of the semi-axes in Table 1, corresponding to 4 kpc for $h = 0.72$) are constrained by the values measured for the galaxy. This model does not fit well the lens configuration ($\chi^2 \sim 34.4$), mainly due to the small uncertainty on the lens galaxy position. When we include the time delays we find a good fit ($\chi^2 \simeq 0.01$) as long as we allow G1 to be misaligned with respect to the observed galaxy. As expected from the reduced surface density compared to the SIE model (Kochanek 2002), we find a much higher value for the Hubble parameter, $h = 0.92$.

5.4. NFW models

We use an NFW model with a fixed break radius of $r_{\rm s} = 10''$ (40 kpc), where the break radius is related to the virial radius through the concentration $c = R_{\rm vir}/r_{\rm s}$. Since $r_{\rm s}$ lies well outside the Einstein radius of the lens, its particular value is not important. This model is not realistic by itself because the shallow $\rho \propto 1/r$ central density cusp of the model will lead to a visible central image. We again find that we can fit the astrometry well even when the position angle of the lens is constrained, but we cannot do so after including the time delays unless we allow the model of G1 to be misaligned relative to the light. In any case, this model leads to a fifth image about 3 mag fainter than A that should be visible on our NICMOS data. This NFW model has a higher surface density near the Einstein ring than the SIE model, so we find a lower value for the Hubble parameter of $h \simeq 0.29$. Using an unphysically small break radius of $r_s = 1''$ raises the density and hence the Hubble parameter to $h \simeq 0.63$.

5.5. De Vaucouleurs plus NFW models

As our final, and most realistic, parametrized model we combine a dVC model constrained by the visible galaxy with an NFW model for the halo. The two components are first constrained to have the same position and position angle, the parameters of the dVC model are constrained by the observations, and the NFW model has a fixed $r_s = 10''$ break radius. This model leads to a poor fit, with $\chi^2 = 6.33$ for $N_{\rm d.o.f.} = 1$. When we free the PA of the NFW model, we find an acceptable fit for $N_{\rm d.o.f.} = 0$, but the misalignment of the NFW model relative to the optical is 40°.

We can further constrain the model by including the three MgII emission line flux ratios from Morgan et al. (2004). We use the line flux ratios instead of those obtained from the light curves, because they should be insensitive to microlensing. With these three additional constraints we still find that the 89.8° position angle of the NFW model is strongly misaligned from the 26.4° position angle of the dVC model, indicating a twisting of the mass isocontours. The model has a reduced χ^2 of 3.2 for $N_{\rm d.o.f.} = 3$. The model flux ratios are significantly different from the constraints. We find $F_A/F_B = 2.81$, $F_A/F_C = 5.01$, and $F_{A1}F_{A2} = 1.26$ while Morgan et al. (2004) report $F_A/F_B = 2.55 \pm 0.60$, $F_A/F_C = 2.02 \pm 0.35$, and $F_{A1}/F_{A2} = 1.61 \pm 0.35$. The match of the flux ratios is better if we do not include the constraints from the time delays. In all cases, F_B/F_C is the most "anomalous" flux ratio, as also found by Morgan et al. (2004). While the differences between the line and continuum flux ratios suggests the presence of long-term microlensing, we see no evidence for the time variability in the flux ratios expected from microlensing. We also note that the broad line flux ratios vary with wavelength (Morgan et al. 2004), which suggests that dust extinction may as well be affecting the flux ratios.

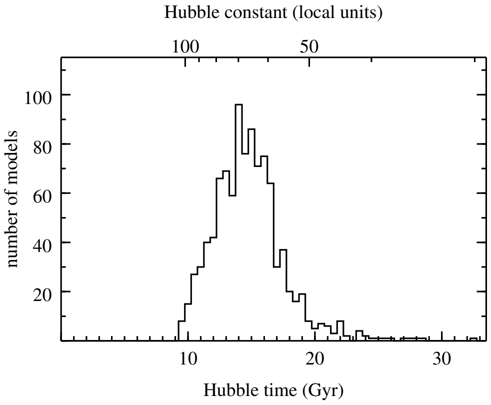


Fig. 11. Distribution of H_0 from 1000 non-parametric models. The bottom-axis gives the Hubble time, the top-axis H_0 in km s^{-1} Mpc^{-1}.

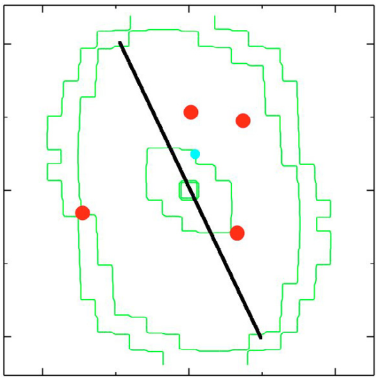


Fig. 12. Mean mass distribution from 1000 pixellated lens models. The red dots are the quasar images and the blue dot the source position. The thick solid line indicates the observed major axis of the lensing galaxy. Each tick-mark measures 1″. The third contour from the outside traces the critical mass density $\Sigma_{\rm crit} = 1.19 \times 10^{11}\ M_\odot$ arcsec^{-2}, and the others are drawn logarithmically from the critical one (each contour is 2.5 larger/smaller than the previous one). North is to the top and East to the left.

6. Non-parametric modeling

We use the non-parametric `PixeLens` (Saha & Williams 2004) method as our second probe of the mass distribution. This approach has the advantage that it makes fewer assumptions about the shape of the G1 than the ellipsoidal parametric models. The models include priors on the steepness of the radial mass profile, imposes smoothness criteria on the profile and we restrict to models symmetric about the lens center. We include two point masses to represent G2 and the group. We run 1000 trial models at a resolution of ~0.23″/pixel which are constrained to fit the image positions and the time delays exactly. For each model we vary the Einstein radii of G2 and the group over the ranges $0.03'' < R_{\rm E}({\rm G2}) < 3''$ and $0.3'' < R_{\rm E}({\rm group}) < 5''$, respectively. Apart from the inclusion of these additional point masses, the method and priors are as explained in detail in Coles (2008). A test, where the technique is used to infer H_0 from a N-body and hydro simulated lens, and an additional discussion of the priors are described in Read et al. (2007). Figure 11 shows the resulting probability distribution for H_0 from the 1000 models. The median value of the distribution is

$$H_0 = 67.1\ ^{+13.0}_{-9.9}\ \ {\rm km\,s^{-1}\,Mpc^{-1}} \tag{6}$$

where the error bars are at 68% confidence. As already illustrated by our parametric modeling, the predicted H_0 value depends on the density gradient of the models. Figure 12 shows the mean surface density contours of the models, and we see a twisting of the contours away from that of the visible galaxy in the outer regions.

7. Conclusions

By combining data from COSMOGRAIL and the SMARTS 1.3 m telescope we measure two independent time delays in the quadruply imaged quasar WFI J2033−4723 (Morgan et al. 2004). The fractional uncertainties of ~4% are among the best obtained so far from an optical monitoring. We obtain the light curves of the quasar images using the MCS deconvolution photometry algorithm (Magain et al. 1998) and then measure time delays using three different approaches with a final estimate of

$$\begin{aligned} \Delta t_{B-A} &= 35.5 \pm 1.4\ \ \text{days}\ \ (3.8\%) \\ \Delta t_{B-C} &= 62.6\ ^{+4.1}_{-2.3}\ \ \text{days}\ \ (^{+6.5\%}_{-3.7\%}) \end{aligned} \tag{7}$$

where A is the mean light curve of the blended of quasar images A_1 and A_2. We find little evidence of microlensing in this system, which makes WFI J2033−4723 a very good system for measuring clean time delays.

The parametric models are consistent with concordance value of H_0 when the lens galaxy has an isothermal mass profile out to the radius of the Einstein ring. As expected, the models allow higher (lower) values as the mass distribution is more centrally concentrated (extended) using de Vaucouleurs (NFW) mass distribution. The non-parametric models predict $H_0 = 67.1^{+13.0}_{-9.9}$ km s^{-1} Mpc^{-1}.

The addition of the time delays as a constraint on the lens models does not alter the mismatch between the observed and predicted image flux ratios. The largest flux ratio anomaly is the 45% difference between the MgII flux ratios found for images B/C. Morgan et al. (2004) also noted that the F_B/F_C flux ratio varies with wavelength, suggesting the presence of chromatic microlensing. The lack of significant variability in the flux ratio over our three year monitoring period suggests either that the effective source velocities in this lens are very low or that the affected images lie in one of the broad demagnified valleys typical of microlensing magnification patterns for low stellar surface densities.

Several galaxies close to the line of sight have a significant impact on the mass modeling. We generally model the potential as the sum of a main lensing galaxy G1, a companion galaxy G2 (∼4″ West of G1), and a nearby group (∼9″ North of G1). Both the parametric and non-parametric models suggest that the isodensity contours of G1 itself must be twisted, with some evidence that the outer regions are aligned with the tidal field of the group rather than the luminous galaxy. This could indicate that G1 is a satellite rather than the central galaxy of the group (e.g. Kuhlen et al. 2007). The twisting seems to be required even though the angular structure of the potential can be adjusted through the companion galaxy G2, an external tidal shear, and in some cases a group halo. Clarifying this issue requires more constraints such as detailed imaging of the Einstein ring image of the quasar host, measuring the redshifts of the nearby galaxies, and measuring the velocity dispersion of G1.

Acknowledgements. We would like to thank all the observers at the EULER telescope for the acquisition of these monitoring data. We also thank Profs. G. Burki and M. Mayor for their help in the flexible scheduling of the observing runs, improving the necessary regular temporal sampling. We are very grateful to Prof. A. Blecha and the technical team of the Geneva Observatory for their efficient help with the electronics of the CCD camera. Many thanks also to Sandrine Sohy for her commitment in the programming part of the deconvolution techniques. Finally, we would like to thank Chuck Keeton for his advice on the use of `LENSMODEL`. This work is supported by ESA and the Belgian Federal Science Policy Office under contract PRODEX 90195. C.S.K. is funded by National Science Foundation grant AST-0708082. COSMOGRAIL is financially supported by the Swiss National Science Foundation (SNSF).

References

Burud, I., Magain, P., Sohy, S., & Hjorth, J. 2001, A&A, 380, 805
Burud, I., Courbin, F., Magain, P., et al. 2002a, A&A, 383, 71
Burud, I., Hjorth, J., Courbin, F., et al. 2002b, A&A, 391, 481
Chantry, V., & Magain, P. 2007, A&A, 470, 467
Coles, J. 2008, ApJ, 679, 17
Courbin, F., Lidman, C., Frye, B. L., et al. 1998, ApJ, 499, L119
DePoy, D. L., Atwood, B., Belville, S. R., et al. 2003, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, ed. M. Iye, & A. F. M. Moorwood, Proc. SPIE, 4841, 827
Eigenbrod, A., Courbin, F., Vuissoz, C., et al. 2005, A&A, 436, 25
Eigenbrod, A., Courbin, F., Meylan, G., Vuissoz, C., & Magain, P. 2006, A&A, 451, 759
Falco, E. E., Impey, C. D., Kochanek, C. S., et al. 1999, ApJ, 523, 617
Ferreras, I., Saha, P., & Burles, S. 2008, MNRAS, 383, 857
Hjorth, J., Burud, I., Jaunsen, A. O., et al. 2002, ApJ, 572, L11
Jean, C., & Surdej, J. 1998, A&A, 339, 729
Keeton, C. R. 2001, ArXiv Astrophysics e-prints
Keeton, C. R., Kochanek, C. S., & Seljak, U. 1997, ApJ, 482, 604
Kochanek, C. S. 2002, ApJ, 578, 25
Kochanek, C. S., Morgan, N. D., Falco, E. E., et al. 2006, ApJ, 640, 47
Krist, J., & Hook, R. 2004, The Tiny Tim's user Guide v.6.3
Kuhlen, M., Diemand, J., & Madau, P. 2007, ApJ, 671, 1135
Lehár, J., Falco, E. E., Kochanek, C. S., et al. 2000, ApJ, 536, 584
Magain, P., Courbin, F., & Sohy, S. 1998, ApJ, 494, 472
Magain, P., Courbin, F., Gillon, M., et al. 2007, A&A, 461, 373
Mao, S., & Schneider, P. 1998, MNRAS, 295, 587
Morgan, C. W., Eyler, M. E., Kochanek, C. S., et al. 2008, ApJ, 676, 80
Morgan, N. D., Caldwell, J. A. R., Schechter, P. L., et al. 2004, AJ, 127, 2617
Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
Oguri, M. 2007, ApJ, 660, 1
Pelt, J., Kayser, R., Refsdal, S., & Schramm, T. 1996, A&A, 305, 97
Poindexter, S., Morgan, N., Kochanek, C. S., & Falco, E. E. 2007, ApJ, 660, 146
Read, J. I., Saha, P., & Macciò, A. V. 2007, ApJ, 667, 645
Refsdal, S. 1964, MNRAS, 128, 307
Saha, P., & Williams, L. L. R. 2004, AJ, 127, 2604
Saha, P., Coles, J., Macciò, A. V., & Williams, L. L. R. 2006a, ApJ, 650, L17
Saha, P., Courbin, F., Sluse, D., Dye, S., & Meylan, G. 2006b, A&A, 450, 461
van Dokkum, P. G. 2001, PASP, 113, 1420
Vuissoz, C., Courbin, F., Sluse, D., et al. 2007, A&A, 464, 845
Yoo, J., Kochanek, C. S., Falco, E. E., & McLeod, B. A. 2006, ApJ, 642, 22

6 Curriculum Vitae

Jonathan Coles

Date of birth : 17. August 1981
Place of birth : Reading, England
Nationality : English
Languages : English (native), German (fluent)

Education

October 2005 – Present	University of Zürich; Zürich, Switzerland Thesis: *The Age of the Universe and the Mass of Stars by Gravitational Lensing* Supervisors: George Lake, Prasenjit Saha
September 2003 – May 2005	Rochester Institute of Technology; Rochester, NY, USA Master of Science, Computer Science Thesis: *Algorithms for Bounding Folkman Numbers* Supervisor: Stanisław Radziszowski
October 2002 – February 2003	University of Osnabrück; Osnabrück, Germany Cognitive Science; Study abroad program
September 1999 – May 2005	Rochester Institute of Technology; Rochester, NY, USA Bachelor of Science, Computer Science with Honors Minor in Mathematics, German Language
September 1995 – May 1999	Sharon High School; Sharon, MA, USA

Publications

1. Weak microlensing (2010)
2. A New Estimate of the Hubble Time with Improved Modeling of Gravitational Lenses (2008)
3. COSMOGRAIL: the COSmological MOnitoring of GRAvItational Lenses. VII. Time delays and the Hubble constant from WFI J2033-4723 (2008)
4. The Hubble Time Inferred from 10 Time Delay Lenses (2006)
5. Computing the Vertex Folkman Number $F_v(2,2,3;4)$ (2004)

Work

July 2004 – October 2004	Optical Tape Systems, Inc.; Shrewsbury, MA, USA Lead software engineer
January 2003 – July 2003	Intevation GmbH; Osnabrück, Germany Software developer
June 2002 – September 2002	Sun Microsystems; Santa Clara, CA, USA Software intern
June 2001 – August 2001	Sun Microsystems; Cupertino, CA, USA Software intern
June 2000 – August 2000	Sun Microsystems; Burlington, MA, USA Software intern
July 1999 – August 1999	Empirix; Waltham, MA, USA Software developer